THE ARCHITECTURE OF METAL COORDINATION GROUPS IN PROTEINS

Tables associated with paper being submitted to Acta Cryst. D (August 2003, revised Dec 2003)

.

Table 1 D Lists of composition, conformation, etc. of all coordination groups in 30% cull set of metalloproteins.)

.

. (Deposited Tables 2D, 4D and Figure 5D, will be deposited with paper, and are not yet available here (Dec 03), but the web Tables 2W and below are available.)

. Table 2 D Statistics about compositioon, chelate loop sizes, common chelate loops, for all metals Ca.. K

Table 4 D Commonly occurring chelate loops with Ca or Zn: numbers found, conformation(s), standard deviations of fi,psi angles in the set, similarity of folds and of local conformations, and name of example of each.

Fig 5 D Details of close approaches of metal atoms, for Zn...Zn up to 6A, and for Ca...Ca up to 7A.

- -

Table 2 W Statistics for all chelate loops in 30% cull set of metalloproteins, of different donor groups with different *seqdif* values; also full statistics for coordination numbers, number of protein donors, no of non protein donors, etc.

. . . statistics . . . . Ca . . . . Mg . . . . Mn . . . . Fe* . . . . Cu . . . . Zn . . . . Na . . . . K . . . .

. . . . . . * Fe coordintaion groups which do not involve haem groups or Fe/S clusters.

.

Table 3 W Details of sequences, conformations, etc in individual Ca and Zn coordination groups which occur more than once in 30% cull set of metalloproteins

.

Table 4 W Details of sequence, conformation, etc. in individual Ca and Zn chelate loops which occur frequently within the 30% cull set of metalloproteins.

.

Table 5 W Proteins with more than one metal per protein chain

Table 5 W (continued) Metal coordination groups where one metal ion is coordinated to more than one protein chain

.

METAL COORDINATION GROUPS IN PROTEINS - Table 1 D

explanation of tables

For each metal there are two tables. The basic information on composition appears in both. The first then includes reliability, nature of protein, etc., while the name of the metal atom and details of conformation are in the second.

The full tables can be downloaded from the web (or from the deposited part of the publication) and manipulated, for example with Micorsoft Excel.

In the first part, nsp is nspan, np is the number of donors from the protein chain, nw the number of water molecules, nn the number of non-protein donor groups, dons are the amino-acid donor groups in the order in which they occur in polypeptide chain, using normal one letter codes for amino-acids, and O for main chain carbonyl oxygen; sd1 to sd7 are the seqdifs (-99 signifies donors are from two different polypeptide chains, -1 is given when the second donor is water or other non-amino-acid donor); his indicates for each f donor groups, including water molecules and small molecule ligands, treating carboxylate always as one group (coordination number,as it would be defined by chemist, is then number of donor groups + number of bidentate carboxylate groups). cn2 is the change in coordination number if sphere is extended from axis - if the number is < 20 then the metal atom is not near a rotation; - if 20- 29 then 20 has been added, signifying that a 2-fold axis is possible; - if 30-39 then 30 has been added, signifying that a 3-fold axis is possible; to establish with certainty it is necessary to look at the spacegroup and coordinates. rms is the r.m.s. deviation of metal to donor atom distances within the coordination sphere from target distances - a useful indicator of quality (0 is good, 0.5 is poor). res is the resolution (A) of the structure determination. carbi indicates bidentate carboxylate groups, e.g. ..b. indicates that the third of four donor groups appears to be a bidentate carboxylate. othdon indicates the type of other donor groups present; Z is a water molecule, U, B, X indicate O, N, S donors in non-protein (small) molecules or ions; J indicates any other non-protein donor group, usually another metal atom. ecno is the E.C. enzyme number, when it is given in the PDB file, followed by the header name given in the PDB file.

In the second part the columns are the same, up to sd7. This is followed by the PDB names of the metal atom and the first donor residue, and then a string indicating the conformation according to the Efimov type categories starting two amino-acid residues before the first donor group (i.e. at relseq = -2); when nspan <30 the conformation sequence is given in full, otherwise it includes the donor residues plus two amino-acids on either side, with a hyphen to indicate where there is a break; at the end are two amino-acid conformations beyond the last donor residue.

Table 1D for Ca : composition,etc . . . . Ca : conformations

Table 1D for Mg : composition,etc . . . . Mg : conformations . . . .

Table 1D for Mn : composition,etc . . . . Mn : conformations . . . .

Table 1D for Fe*: composition,etc . . . . Fe*: conformations . . . .

Table 1D for Cu : composition,etc . . . . Cu : conformations . . . .

Table 1D for Zn : composition,etc . . . . Zn : conformations . . . .

Table 1D for Na : composition,etc . . . . Na : conformations . . . .

Table 1D for K : composition,etc . . . . K : conformations . . . .

*Fe coordination groups excluding those which involve haem groups and Fe/S clusters.

```
no  cngpname    nsp  np  nw  nn  dons     met  sd1  sd2  sd3  sd4  sd5  sd6  sd7  his   cn  pdb  cn2  rms  res   carbi    othdon   ecno
  1 1scf_D   54     4   2   2   0  DD        Ca    4   -1   -1   -1   -1   -1   -1 ..     4 1scf   0 0.4  2.2   .b       | ZZ                   HOR  MONE/GROWTH F
  2 1rdr_   233    96   2   5   0  DD        Ca   96   -1   -1   -1   -1   -1   -1 ..     7 1rdr   0  0.3  2.4   ..      | ZZZZZ  2.7.7.48;      NUC    LEOTIDYLTRANS
  3 1fsu_    53   248   4   0   3  DD    DN  Ca    1   -1   -1   -1  246    1   -1 ..    .7 1fsu   0  0.3  2.5   ..    b. | UXU    3.1.6.12;     HYD        ROLASE
  4 1bn8_A  184    43   3   1   0  DDD       Ca   39    4   -1   -1   -1   -1   -1 ...    4 1bn8   1  0.5  1.8   b..     |  Z     4.2.2.2        LYA   SE
  5 1alv_B  135    91   4   2   0  DDDN      Ca   88    2    1   -1   -1   -1   -1 ....   6 1alv   0  0.1  1.9   .bb.    |   ZZ   3.4.22.17;    CAL        CIUM  BINDING
  6 2scp_A   16    11   5   1   0  DDDOD     Ca    2    2    2    5   -1   -1   -1 ..... 6 2scp   0  0.1  2.0   ....b   |    Z                   BIN DING  PROTEIN
  7 1cdl_A   20    11   5   0   0  DDDOE     Ca    2    2    2    5   -1   -1   -1 ..... 5 1cdl   0  0.4  2.2   .b..b   |                        CAL CIUM-BINDING
  8 1cdl_B  129    11   5   0   0  DDDOE     Ca    2    2    2    5   -1   -1   -1 ..... 5 1cdl   0  0.4  2.2   .b..b   |                        CAL CIUM-BINDING
  9 1acc_   177    11   5   1   0  DDDOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 1acc   1  0.2  2.1   .....   |    Z                   TOX IN
 10 1g4y_R   20    11   5   1   0  DDDOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 1g4y   0  0.2  1.6   ....b   |    Z                   SIG NALING  PROTEI
 11 1sra_   257    11   5   1   0  DDDOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 1sra   0  0.2  2.0   ....b   |    Z                   CAL CIUM-BINDING
 12 1vrk_A   20    11   5   1   0  DDDOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 1vrk   0  0.1  1.9   ....b   |    Z   2.7.1.117;    COM        PLEX(CALCIUM-
 13 1vrk_A  129    11   5   1   0  DDDOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 1vrk   0  0.2  1.9   ....b   |    Z   2.7.1.117;    COM         PLEX(CALCIUM-
 14 2pvb_A   90    11   5   1   0  DDDOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 2pvb   0  0.1  0.9   ....b   |    Z                   METAL  BINDING  PR
 15 1acc_   179    56   6   0   0  DDEOOD    Ca    2    7   34    3   10   -1   -1 ......6 1acc   0  0.3  2.1   ..b...  |                       TOX IN
 16 2por_    93     8   4   2   0  DDND      Ca    2    5    1   -1   -1   -1   -1 ....   6 2por   0  0.2  1.8   bb..    |   ZZ                   INT EGRAL  MEMBRAN
 17 1cdl_B   56    11   6   0   0  DDNNOE    Ca    2    2    0    2    5   -1   -1 ......6 1cdl   0  0.6  2.2   .....b  |                        CAL CIUM-BINDING
 18 1cdl_C   56    11   6   0   0  DDNODE    Ca    2    2    2    2    3   -1   -1 ......6 1cdl   0  0.6  2.2   .....b  |                        CAL CIUM-BINDING
 19 1cdl_A   93    11   5   0   0  DDNOE     Ca    2    2    2    5   -1   -1   -1 ..... 5 1cdl   0  0.3  2.2   .b..b   |                        CAL CIUM-BINDING
 20 1cdl_A   56    11   5   0   0  DDNOE     Ca    2    2    2    5   -1   -1   -1 ..... 5 1cdl   1  0.4  2.2   ....b   |                        CAL CIUM-BINDING
 21 1g4y_R   56    11   5   0   0  DDNOE     Ca    2    2    2    5   -1   -1   -1 ..... 5 1g4y   0  0.4  1.6   ....b   |                        SIG NALING  PROTEI
 22 1rec_   110    11   5   1   0  DDNOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 1rec   0  0.2  1.9   ....b   |    Z                   CAL CIUM-BINDING
 23 1vrk_A   56    11   5   1   0  DDNOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 1vrk   0  0.1  1.9   ....b   |    Z   2.7.1.117;    COM        PLEX(CALCIUM-
 24 1fzc_B  381     4   3   2   0  DDO       Ca    2    2   -1   -1   -1   -1   -1 ...    5 1fzc   0  0.2  2.3   b..     |  ZZ                    BLO OD COAGULATIO
 25 2sns_    21    20   3   0   1  DDO       Ca   19    1   -1   -1   -1   -1   -1 ...    4 2sns   0  0.3  1.5   ...     |  U     3.1.33.1)      HYD  ROLASE  (PHOSP
 26 1kit_   621    62   3   1   0  DDO       Ca   61    1   -1   -1   -1   -1   -1 ...    4 1kit   0  0.2  2.3   bb.     |  Z     3.2.1.18;      HYD  ROLASE
 27 2mas_B   10   232   4   1   2  DDOD      Ca    5  111  116   -1   -1   -1   -1 ....   7 2mas   0  0.2  2.3   .b..    |   UUZ  3.2.2.1;      HYD                ROLASE
 28 1fzc_C  318     6   4   0   0  DDOO      Ca    2    2    2   -1   -1   -1   -1 ....   4 1fzc   0  0.2  2.3   b...    |                       BLO OD COAGULATIO
 29 3fib_   318     6   4   2   0  DDOO      Ca    2    2    2   -1   -1   -1   -1 ....   6 3fib   0  0.2  2.1   b...    |   ZZ                  BLO OD COAGULATIO
 30 1qge_E  241    54   4   2   0  DDOO      Ca   46    4    4   -1   -1   -1   -1 ....   6 1qge   0  0.2  1.7   ....    |   ZZ   3.1.1.3;       HYD       ROLASE
 31 1ava_A  127    21   5   1   0  DDOOD     Ca   15    1    3    2   -1   -1   -1 ..... 6 1ava   1  0.2  1.9   .....   |    Z   3.2.1.1;       HYD       ROLASE  INHIBI
```

```
32 1alv_B 180    11   5   1   0 DDSOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 1alv  0 0.1 1.9  ....b   |     Z  3.4.22.17;      CAL      CIUM BINDING
33 2pvb_A  51    11   6   0   0 DDSOEE    Ca    2    2    2    2    3   -1   -1 ......6 2pvb  0 0.1 0.9  .....b  |                              METAL BINDING PR
34 1alv_A 150    11   5   1   0 DDTOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 1alv  0 0.2 1.9  ....b   |     Z  3.4.22.17;      CAL      CIUM BINDING
35 1f7l_A   8    50   2   3   0 DE          Ca   50   -1   -1   -1   -1   -1   -1 ..     5 1f7l  0 0.2 1.5  b.       | ZZZ   2.7.8.7;         TRANSFERASE
36 2msb_A 161    33   5   1   0 DEDOD     Ca    4   23    5    1   -1   -1   -1 ..... 6 2msb  0 0.2 1.7  bb...   |     Z                         LEC   TIN
37 1hyo_A 126   107   4   0   2 DEED      Ca   73    2   32   -1   -1   -1   -1 ....  6 1hyo  0 0.1 1.3  ....    |   UU  3.7.1.2;         HYD          ROLASE
38 1ezm_  136    49   5   1   0 DEEDO     Ca   36    3    8    2   -1   -1   -1 ..... 6 1ezm  0 0.3 1.5  ..b..   |     Z  3.4.24.26)      HYD          ROLASE
39 1tn3_  116    35   5   1   0 DEOON     Ca    4   27    3    1   -1   -1   -1 ..... 6 1tn3  0 0.2 2.0  bb...   |     Z                         LECTIN
40 2sas_   19    11   5   0   0 DNDOD     Ca    2    2    2    5   -1   -1   -1 ..... 5 2sas  0 0.2 2.4  ....b   |                            CALCIUM-BINDING
41 1a2x_A 139    11   5   0   0 DNDOE     Ca    2    2    2    5   -1   -1   -1 ..... 5 1a2x  1 0.4 2.3  ..b.b   |                              COM PLEX (SKELETA
42 1a2x_A 103    11   5   0   0 DNDOE     Ca    2    2    2    5   -1   -1   -1 ..... 5 1a2x  2 0.2 2.3  ....b   |                              COMPLEX (SKELETA
43 2scp_B 104    11   5   0   0 DNDOE     Ca    2    2    2    5   -1   -1   -1 ..... 5 2scp  0 0.3 2.0  ....b   |                              BINDING PROTEIN
44 2sas_   70    11   5   1   0 DNDOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 2sas  0 0.4 2.4  ....b   |     Z                       CALCIUM-BINDING
45 2scp_B 138    11   5   1   0 DNDOE     Ca    2    2    2    5   -1   -1   -1 ..... 6 2scp  0 0.1 2.0  ....b   |     Z                        BINDING PROTEIN
46 1gca_  134    71   6   0   0 DNDOQE    Ca    2    2    2    2   63   -1   -1 ......6 1gca  0 0.1 1.7  .....b  |                            GALACTOSE-BINDIN
47 1i8a_A  81    13   5   1   0 DNEDD     Ca    2    8    2    1   -1   -1   -1 ..... 6 1i8a  0 0.3 1.9  ...bb   |     Z  3.2.1.8;        HYD     ROLASE
48 1sac_C  58    80   5   0   0 DNEOD     Ca    1   77    1    1   -1   -1   -1 ..... 5 1sac  0 0.2 2.0  b.b..   |                              AMYLOID PROTEIN
49 1sac_A  58   242   6   0   0 DNEODO    Ca    1   77    1    1  162   -1   -1 ......6 1sac  0 0.3 2.0  b.b...  |                           AMY     LOID PROTEIN
50 1bfd_  428    29   3   1   2 DNO        Ca   27    2   -1   -1   -1   -1   -1 ...    6 1bfd  0 0.2 1.6  ...      |  UUZ  4.1.1.7;        LYA    SE
51 1trk_A 157    32   3   1   2 DNO       Ca   30    2   -1   -1   -1   -1   -1 ...    6 1trk  0 0.2 2.0  ...      |  UUZ  2.2.1.1)        TRA    NSFERASE(KETO
52 1b0p_B 983    78   5   0   0 DNOOO     Ca    2   71    3    2   -1   -1   -1 ..... 5 1b0p  1 0.4 2.3  .....   |        1.2.7.1;        OXI    DOREDUCTASE
53 1qq9_A   3   263   4   2   0 DODD      Ca    1  258    4   -1   -1   -1   -1 ....  6 1qq9  0 0.2 1.5  ...b    |   ZZ  3.4.11.-         HYD         ROLASE
54 1e43_A 159    45   5   1   0 DODDD     Ca   22    2   19    2   -1   -1   -1 ..... 6 1e43  0 0.1 1.7  b....   |     Z  3.2.1.1;        HYD          ROLASE
55 1oac_A 533   146   5   1   0 DODDO     Ca    1    1  143    1   -1   -1   -1 ..... 6 1oac  0 0.1 2.0  .....   |     Z  1.4.3.6         OXI          DOREDUCTASE
56 1i8a_A  60    95   5   1   0 DODDO     Ca    2   12   80    1   -1   -1   -1 ..... 6 1i8a  0 0.3 1.9  b.b..   |     Z  3.2.1.8;        HYD        ROLASE
57 1qho_A  76    26   5   1   0 DODEE     Ca    1    2   22    1   -1   -1   -1 ..... 6 1qho  1 0.2 1.7  ....b   |     Z  3.2.1.133;      HYD          ROLASE
58 1sra_  222    12   5   1   0 DODOE     Ca    3    2    2    5   -1   -1   -1 ..... 6 1sra  0 0.2 2.0  ....b   |     Z                        CALCIUM-BINDING
59 1cvr_A  78   180   3   3   0 DOE       Ca  171    9   -1   -1   -1   -1   -1 ...    6 1cvr  0 0.2 2.0  b.b      |  ZZZ  3.4.22.37;      HYD            ROLASE
60 1dx5_J 423    20   6   1   0 DOENOO    Ca    1    2   13    1    3   -1   -1 ......7 1dx5  0 0.2 2.3  b.....  |     Z 3.4.21.5;       SER            INE PROTEINAS
61 1nls_   10     9   4   2   0 DOND       Ca    2    2    5   -1   -1   -1   -1 ....  6 1nls  0 0.1 0.9  b...     |   ZZ                        AGGLUTIN
62 1qho_A  21    29   6   1   0 DONNOD    Ca    2    3    1   21    2   -1   -1 ......7 1qho  0 0.2 1.7  ......  |     Z 3.2.1.133;      HYD            ROLASE
63 1ee6_A  80    23   3   3   0 DOO        Ca    1   22   -1   -1   -1   -1   -1 ...    6 1ee6  0 0.3 2.3  ...      |  ZZZ  4.2.2.2         LYA    SE
64 1d2s_A  50   110   3   3   0 DOO       Ca    2  108   -1   -1   -1   -1   -1 ...    6 1d2s  0 0.1 1.5  b..      |  ZZZ                        TRA  NSPORT PROTEI
```

```
65 1svy_   222      7   3   3   0 DOO         Ca    5    2   -1   -1   -1   -1   -1 ...    6 1svy   0 0.3 1.8   ...       |  ZZZ                      ACTIN-BINDING PR
66 1dyk_A2808    68   4   1   0 DOOD       Ca   17   49    2   -1   -1   -1   -1 ....   5 1dyk   0 0.2 2.0   ....      |    Z                        MET  AL BINDING PR
67 1dyk_A2982    73   4   0   0 DOOD       Ca   17   54    2   -1   -1   -1   -1 ....   4 1dyk   0 0.2 2.0   ....      |                             MET AL  BINDING PR
68 1fjs_A   70    10   5   1   0 DOOEE      Ca    2    3    2    3   -1   -1   -1 ..... 6 1fjs   0 0.3 1.9   .....     |     Z   3.4.21.6;        BLO  OD CLOTTING
69 1gcy_A    1    16   6   0   0 DOOHDE     Ca    1   11    0    3    1   -1   -1 .. d..6 1gcy   0 0.2 1.6   b.....    |         3.2.1.60;        HYD        ROLASE
70 1ga6_A 328    20   5   1   0 DOOOD      Ca    1   15    2    2   -1   -1   -1 ..... 6 1ga6   0 0.1 1.0   .....     |     Z   3.4.23.37;       HYD          ROLASE
71 1i76_A 154    26   6   0   0 DOOODE     Ca    1    2    2   18    3   -1   -1 ......6 1i76   0 0.1 1.2   ......   |         3.4.24.34;       HYD         ROLASE
72 1hdf_B   45    47   4   2   0 DOOS       Ca   17   28    2   -1   -1   -1   -1 ....   6 1hdf   0 0.3 2.3   ....      |    ZZ                       STRUCTURAL PROTE
73 1edm_C   47    18   5   0   0 DOQDO      Ca    1    2   14    1   -1   -1   -1 ..... 5 1edm   0 0.2 1.5   ...b.     |                             COA GULATION FACT
74 1kap_P 446     8   5   1   0 DSDOD      Ca    2    2    2    2   -1   -1   -1 ... . 6 1kap   0 0.3 1.6   ....b     |     Z   3.4.24.-;        ZIN   C METALLOPROT
75 2sas_   115    11   5   0   0 DSDOE      Ca    2    2    2    5   -1   -1   -1 ..... 5 2sas   0 0.2 2.4   ....b     |                           CALCIUM-BINDING
76 1c7k_A   76      2   2   4   0 DT         Ca    2   -1   -1   -1   -1   -1   -1 ..     6 1c7k   0 0.1 1.0   b.         | ZZZZ   3.4.24.-         HYD   ROLASE
77 1hei_A 441      4   2   1   0 DT         Ca    4   -1   -1   -1   -1   -1   -1 ..     3 1hei   0 0.4 2.1   b.         | Z                      ICASE HEL
78 2cbl_A 229    11   5   1   0 DTNOE      Ca    2    2    2    5   -1   -1   -1 ..... 6 2cbl   0 0.1 2.1   ....b     |     Z                        COMPLEX (PROTO-O
79 1d0b_A   55      4   2   3   0 ED         Ca    4   -1   -1   -1   -1   -1   -1 ..     5 1d0b   0 0.3 1.9   .b         | ZZZ                          CELL ADHESION
80 2por_    80    28   2   4   0 ED         Ca   28   -1   -1   -1   -1   -1   -1 ..     6 2por   0 0.1 1.8   b.         | ZZZZ                        INTEGRAL MEMBRAN
81 2msb_A 165    29   2   4   0 ED         Ca   29   -1   -1   -1   -1   -1   -1 ..     6 2msb   0 0.2 1.7   .b         | ZZZZ                         LEC TIN
82 1g0h_A   65   136   4   0   2 EDDD       Ca   16    3 117   -1   -1   -1   -1 ....   6 1g0h   0 0.3 2.3   ....      |    UU   3.1.3.25;        HYD            ROLASE
83 1cb8_A 405    12   4   2   0 EDDO       Ca    2    9    1   -1   -1   -1   -1 ....   6 1cb8   0 0.2 1.9   .b..      |    ZZ   4.2.2.5          LYA      SE
84 1g0h_B 365    18   3   0   1 EDO        Ca   16    2   -1   -1   -1   -1   -1 ...    4 1g0h   1 0.4 2.3   .b.        |  U     3.1.3.25;        HYD      ROLASE
85 1atl_A    9   191   4   1   0 EDON       Ca   84  104    3   -1   -1   -1   -1 ....   5 1atl   0 0.2 1.8   .b..      |    Z    3.4.24.42;       MET       ALLOENDOPEPTI
86 1sac_C 136    12   3   0   0 EDQ        Ca    2   10   -1   -1   -1   -1   -1 ...    3 1sac   0 0.2 2.0   .b.        |                             AMYLOID PROTEIN
87 1blx_A   69      3   2   4   0 EE         Ca    3   -1   -1   -1   -1   -1   -1 ..     6 1blx   0 0.2 1.9   ..         | ZZZZ                        COMPLEX (INHIBIT
88 1i7o_A 276    31   3   3   0 EED        Ca    2   29   -1   -1   -1   -1   -1 ...    6 1i7o   0 0.1 1.7   ...        |  ZZZ   5.3.3.10,        ISO    MERASE, LYASE
89 1cvr_A 293      9   3   1   0 EEH        Ca    3    6   -1   -1   -1   -1   -1 ..e    4 1cvr   1 0.5 2.0   b..        |   Z     3.4.22.37;       HYD   ROLASE
90 1cvr_A 161   131   3   3   0 EHD        Ca    5 126   -1   -1   -1   -1   -1 .d.    6 1cvr   1 0.3 2.0   ...        | ZZZ   3.4.22.37;       HYD            ROLASE
91 1g72_A 171    84   2   0   3 EN         Ca   84   -1   -1   -1   -1   -1   -1 ..     5 1g72   0 0.3 1.9   b.         | UBU    1.1.99.8;        OXI       DOREDUCTASE
92 1byf_B   86    22   5   0   2 ENDOD      Ca    3   18    1    0   -1   -1   -1 ..... 7 1byf   0 0.1 2.0   .....     |     UU                       SUG AR BINDING PR
93 2msb_A 185    21   6   0   0 ENENOD     Ca    2    6   12    1    0   -1   -1 ......6 2msb   0 0.2 1.7   ......   |                            LEC  TIN
94 1cru_B 253    10   2   4   0 EO         Ca   10   -1   -1   -1   -1   -1   -1 ..     6 1cru   0 0.1 1.5   b.         | ZZZZ   1.1.99.17;       OXI       DOREDUCTASE
95 1bag_   276    37   2   3   0 EO         Ca   37   -1   -1   -1   -1   -1   -1 ..     5 1bag   0 0.2 2.5   b.         | ZZZ    3.2.1.1;         ALP   HA-AMYLASE
96 1oac_A 573    99   4   2   0 EODE       Ca   94    3    2   -1   -1   -1   -1 ....   6 1oac   0 0.2 2.0   b...      |    ZZ   1.4.3.6          OXI      DOREDUCTASE
97 1g5c_F   97    24   3   3   0 EOO        Ca   21    3   -1   -1   -1   -1   -1 ...    6 1g5c   0 0.3 2.1   b..        |  ZZZ   4.2.1.1;         LYA      SE
```

```
 98 1ava_B 108     9   4   1   0 EOOD      Ca    3    2    4  -1  -1  -1  -1 ....  5 1ava  0 0.4 1.9  b..b    |   Z   3.2.1.1;       HYD    ROLASE INHIBI
 99 1sbw_A  70    10   4   2   0 EOOE      Ca    2    3    5  -1  -1  -1  -1 ....  6 1sbw  0 0.2 1.8  ....    |   ZZ  3.4.21.4;      HYD      ROLASE/HYDROL
100 2btc_E  70    10   4   2   0 EOOE      Ca    2    3    5  -1  -1  -1  -1 ....  6 2btc  0 0.2 1.5  ....    |   ZZ  3.4.21.4;      HYD    ROLASE/HYDROL
101 1avw_A  70    10   5   1   0 EOOEE     Ca    2    3    2    3  -1  -1  -1 ..... 6 1avw  0 0.1 1.8  .....   |    Z  3.4.21.4;      COM        PLEX (PROTEIN
102 1fs7_A 217    59   4   2   0 EOOQ      Ca    1   56    2  -1  -1  -1  -1 ....  6 1fs7  0 0.1 1.6  b...    |   ZZ                    OXIDOREDUCTASE
103 1pyt_C 470    10   5   0   0 EOOQE     Ca    2    3    2    3  -1  -1  -1 ..... 5 1pyt  1 0.3 2.3  .....   |       3.4.17.1;      TER    NARY COMPLEX
104 1gcy_A 116    81   5   1   0 NDODO     Ca   35    3    8   35  -1  -1  -1 ..... 6 1gcy  0 0.2 1.6  .b...   |    Z  3.2.1.60;      HYD           ROLASE
105 1ava_A  91    92   5   1   0 NDODO     Ca   47    3    7   35  -1  -1  -1 ..... 6 1ava  0 0.2 1.9  .b...   |    Z  3.2.1.1;       HYD         ROLASE INHIBI
106 2por_  116    24   4   1   0 NDOO      Ca   20    2    2  -1  -1  -1  -1 ....  5 2por  0 0.2 1.8  .b..    |   Z                     INTEGRAL MEMBRAN
107 1jg8_D 330     3   2   1   0 NE        Ca    3  -1  -1  -1  -1  -1  -1 ..     3 1jg8  0 0.1 1.8  ..      | Z      4.1.2.5;       LYASE
108 1e43_A 444     3   2   5   0 NE        Ca    3  -1  -1  -1  -1  -1  -1 ..     7 1e43  0 0.2 1.7  .b      | ZZZZZ 3.2.1.1;       HYD       ROLASE
109 1i9z_A 568    29   2   4   0 NE        Ca   29  -1  -1  -1  -1  -1  -1 ..     6 1i9z  2 0.5 1.8  .b      | ZZZZ                   HYD ROLASE
110 1djx_B 312    78   4   1   1 NEDE      Ca   29    2   47  -1  -1  -1  -1 ....  6 1djx  0 0.2 2.3  ..b.    |   UZ  3.1.4.11;      LIP       ID DEGRADATIO
111 1e29_A  49     1   2   6   0 NN        Ca    1  -1  -1  -1  -1  -1  -1 ..     8 1e29  0 0.1 1.2  ..      | ZZZZZZ                 ELECTRON TRANSPO
112 1egi_A 747     1   3   1   0 NOD       Ca    1    0  -1  -1  -1  -1  -1 ...    4 1egi  0 0.2 2.3  ...     |  Z                   SUGAR BINDING PR
113 1e43_A 102   133   5   1   0 NODDO     Ca   92    0    6   35  -1  -1  -1 ....  6 1e43  1 0.2 1.7  .....   |    Z  3.2.1.1;       HYD          ROLASE
114 1bag_  101    79   4   1   0 NODO      Ca   36    9   34  -1  -1  -1  -1 ...   5 1bag  0 0.2 2.5  ..b.    |   Z   3.2.1.1;       ALP     HA-AMYLASE
115 7taa_  121    89   4   3   0 NODO      Ca   41   13   35  -1  -1  -1  -1 ...   7 7taa  0 0.2 2.0  ..b.    |   ZZZ 3.2.1.1        HYD        ROLASE
116 1clv_A  98    91   4   3   0 NODO      Ca   48    9   34  -1  -1  -1  -1 ...   7 1clv  0 0.2 2.0  ..b.    |   ZZZ 3.2.1.1;       HYD        ROLASE
117 1qho_A 131   101   4   3   0 NODO      Ca   53   14   34  -1  -1  -1  -1 ...   7 1qho  0 0.2 1.7  ..b.    |   ZZZ 3.2.1.133;     HYD                 ROLASE
118 1nps_A  36    43   3   3   0 NOS       Ca   17   26  -1  -1  -1  -1  -1 ...   6 1nps  0 0.5 1.8  ...     |  ZZZ                   SIGNALING PROTEI
119 1e29_A  35     0   2   4   0 OD        Ca    0  -1  -1  -1  -1  -1  -1 ..     6 1e29  0 0.3 1.2  ..      | ZZZZ                   ELECTRON TRANSPO
120 1bag_  169     2   2   3   0 OD        Ca    2  -1  -1  -1  -1  -1  -1 ..     5 1bag  0 0.1 2.5  .b      | ZZZ   3.2.1.1;       ALP HA-AMYLASE
121 1d0b_A  49     2   2   5   0 OD        Ca    2  -1  -1  -1  -1  -1  -1 ..     7 1d0b  0 0.2 1.9  ..      | ZZZZZ                  CELL ADHESION
122 1qh4_D  41     3   2   0   0 OD        Ca    3  -1  -1  -1  -1  -1  -1 ..     2 1qh4  0 0.2 1.4  .b      |        2.7.3.2;       TRANSFERASE
123 3std_C  51     4   2   2   0 OD        Ca    4  -1  -1  -1  -1  -1  -1 ..     4 3std  0 0.3 1.6  ..      | ZZ     4.2.1.94;      LYASE
124 1wdc_C  19     8   6   1   0 ODDODO    Ca    0    3    1    2    2  -1  -1 ......7 1wdc  0 0.2 2.0  ......  |     Z                  MUS  CLE PROTEIN
125 1djx_A 651    26   3   2   0 ODN       Ca    2   24  -1  -1  -1  -1  -1 ...   5 1djx  1 0.4 2.3  .b.     |  ZZ   3.1.4.11;      LIP     ID DEGRADATIO
126 1clc_  520     5   3   3   0 ODO       Ca    3    2  -1  -1  -1  -1  -1 ...   6 1clc  0 0.1 1.9  ...     |  ZZZ  3.2.1.4;       GLY COSYL HYDROLA
127 1e54_A  29   280   3   0   0 ODO       Ca   35  245  -1  -1  -1  -1  -1 ...   3 1e54  0 0.2 2.1  ...     |                        OUTER MEMBRANE P
128 1ajj_   22    14   6   0   0 ODODDE    Ca    3    2    2    6    1  -1  -1 ......6 1ajj  0 0.1 1.7  ......  |                      RECEPTOR
129 1i8a_A  10   120   5   1   0 ODODE     Ca    2    2    2  114  -1  -1  -1 ..... 6 1i8a  0 0.2 1.9  ....b   |    Z  3.2.1.8;       HYD        ROLASE
130 1aru_   57    22   5   2   0 ODODS     Ca    0   18    2    2  -1  -1  -1 ..... 7 1aru  0 0.2 1.6  .....   |    ZZ 1.11.1.7;      PER      OXIDASE (DONO
```

```
131 1kap_P 288    41   4   2   0 ODOE       Ca    2   37    2   -1   -1   -1   -1 ....   6 1kap   0 0.1 1.6   ...b     |    ZZ   3.4.24.-;        ZIN       C METALLOPROT
132 1cvr_A 100     7   4   2   0 ODOE       Ca    3    2    2   -1   -1   -1   -1 ....   6 1cvr   0 0.2 2.0   ...b     |    ZZ   3.4.22.37;       HYD        ROLASE
133 1alv_B 107    10   4   2   0 ODOE       Ca    3    2    5   -1   -1   -1   -1 ....   6 1alv   0 0.1 1.9   ...b     |    ZZ   3.4.22.17;       CAL        CIUM BINDING
134 1tf4_B 504    71   5   1   0 ODOND      Ca    2   65    3    1   -1   -1   -1 ..... 6 1tf4   0 0.1 1.9   .b...    |     Z   3.2.1.4;         GLY       COSYL HYDROLA
135 1pa2_A  43     9   6   1   0 ODOODS     Ca    0    3    2    2    2   -1   -1 ......7 1pa2   0 0.1 1.5   ......   |      Z 1.11.1.7;        OXI       DOREDUCTASE
136 1e8u_A 261    35   5   0   0 ODOOO      Ca    0    3    2   30   -1   -1   -1 ..... 5 1e8u   0 0.1 2.0   .....    |         3.2.1.18         SIA      LIDASE
137 4sgb_E 120   123   5   1   0 ODOOY      Ca    0    1  122    0   -1   -1   -1 ..... 6 4sgb   0 0.1 2.1   .....    |     Z                      COM    PLEX(SERINE P
138 1e8u_B 261    35   6   0   0 ODOSOO     Ca    0    3    0    2   30   -1   -1 ......6 1e8u   0 0.2 2.0   ......   |         3.2.1.18         SIA       LIDASE
139 1g5c_C 147     0   2   3   0 OE          Ca    0   -1   -1   -1   -1   -1   -1 ..     5 1g5c   0 0.4 2.1   ..        | ZZZ    4.2.1.1;         LYA  SE
140 1ayo_B 120     1   2   3   0 OE          Ca    1   -1   -1   -1   -1   -1   -1 ..     5 1ayo   0 0.1 1.9   ..        | ZZZ                          MACROGLOBULIN
141 1c1y_B 123     2   2   2   0 OE          Ca    2   -1   -1   -1   -1   -1   -1 ..     4 1c1y   0 0.2 1.9   .b        | ZZ     2.7.1.-;         SIGNALING PROTEI
142 1qhd_A 151   189   2   2   0 OE          Ca  189   -1   -1   -1   -1   -1   -1 ..     4 1qhd   1 0.2 2.0   ..        | ZZ                         VIRUS/VIRAL PROT
143 1dyo_A  14   135   5   0   0 OEOOD      Ca    2   23    3  107   -1   -1   -1 ..... 5 1dyo   0 0.4 2.1   ....b    |                            CAR    BOHYDRATE-BIN
144 1fmj_A 118     4   3   0   0 OND         Ca    3    1   -1   -1   -1   -1   -1 ...    3 1fmj   0 0.2 2.0   ...       |                           NSFERASE
145 1clc_  236    10   5   1   0 ONODD      Ca    3    2    2    3   -1   -1   -1 ..... 6 1clc   0 0.1 1.9   .....    |     Z   3.2.1.4;         GLY    COSYL HYDROLA
146 1cru_B 247     1   2   2   3 OO          Ca    1   -1   -1   -1   -1   -1   -1 ..     7 1cru   0 0.1 1.5   ..        | UBUZZ 1.1.99.17;       OXI        DOREDUCTASE
147 1c8n_A 173     2   2   0   0 OO          Ca    2   -1   -1   -1   -1   -1   -1 ..     2 1c8n   0 0.5 2.3   ..        |                          US/VIRAL PROT
148 1kap_P 370    30   4   2   0 OODD       Ca    2    2   26   -1   -1   -1   -1 ....   6 1kap   1 0.2 1.6   ....     |    ZZ   3.4.24.-;        ZIN       C METALLOPROT
149 1lpb_B 187     8   4   2   0 OODD       Ca    3    2    3   -1   -1   -1   -1 ....   6 1lpb   0 0.2 2.5   ...b     |    ZZ   3.1.1.3) C       HYD        ROLASE(CARBOX
150 1cru_B 269    40   4   2   0 OODE       Ca    2    2   36   -1   -1   -1   -1 ....   6 1cru   0 0.1 1.5   ...b     |    ZZ   1.1.99.17;       OXI        DOREDUCTASE
151 1f8e_A 293    54   4   2   0 OODO       Ca    4   27   23   -1   -1   -1   -1 ....   6 1f8e   0 0.4 1.4   ....     |    ZZ   3.2.1.18;        HYD          ROLASE/HYDROL
152 1kap_P 361    29   5   1   0 OODOD      Ca    2    2   18    7   -1   -1   -1 ..... 6 1kap   0 0.1 1.6   .....    |     Z   3.4.24.-;        ZIN       C METALLOPROT
153 1kap_P 334    22   6   0   0 OODOOD     Ca    2    2   13    2    3   -1   -1 ......6 1kap   0 0.3 1.6   .....b   |         3.4.24.-;        ZIN        C METALLOPROT
154 1kap_P 352    22   6   0   0 OODOOD     Ca    2    2   13    2    3   -1   -1 ......6 1kap   0 0.3 1.6   .....b   |         3.4.24.-;        ZIN        C METALLOPROT
155 1sra_  241     5   3   2   0 OOE         Ca    2    3   -1   -1   -1   -1   -1 ...    5 1sra   1 0.4 2.0   ...       |   ZZ                      CALCIUM-BINDING
156 1ej8_A 124    44   3   3   0 OON         Ca   44    0   -1   -1   -1   -1   -1 ...    6 1ej8   0 0.2 1.5   ...       |  ZZZ                        CHAPERONE
157 1kit_  253    60   6   0   0 OONDOT     Ca    3    0   33   24    0   -1   -1 ......6 1kit   0 0.2 2.3   ...b..   |         3.2.1.18;        HYD       ROLASE
158 1kap_P 343    22   6   0   0 OONOOD     Ca    2    2   13    2    3   -1   -1 ......6 1kap   0 0.3 1.6   .....b   |         3.4.24.-;        ZIN        C METALLOPROT
159 1dmu_A  99     5   3   3   0 OOO         Ca    2    3   -1   -1   -1   -1   -1 ...    6 1dmu   0 0.2 2.2   ...       |   ZZZ   3.1.21.4;        HYD        ROLASE/DNA
160 1a0t_P 454     8   3   0   0 OOO         Ca    3    5   -1   -1   -1   -1   -1 ...    3 1a0t   0 0.1 2.4   ...       |                           OUTER MEMBRANE P
161 1poc_    8    27   4   0   2 OOOD       Ca    2    2   23   -1   -1   -1   -1 ....   6 1poc   0 0.1 2.0   ...b     |    UU   3.1.1.4) C       HYD        ROLASE
162 1axn_   32    44   4   0   0 OOOD       Ca    2    2   40   -1   -1   -1   -1 ....   4 1axn   0 0.4 1.8   ...b     |                            CALCIUM/PHOSPHOL
163 2sic_E 169    28   4   2   0 OOOD       Ca    5   21    2   -1   -1   -1   -1 ....   6 2sic   1 0.6 1.8   ....     |    ZZ   3.4.21.14)       COM         PLEX (PROTEIN
```

```
164 1i76_A 137    36   4   2   0  OOOD       Ca   32    2    2   -1   -1   -1   -1 ....   6 1i76   0 0.1 1.2   ....     |    ZZ   3.4.24.34;       HYD          ROLASE
165 1e43_A 300   130   5   1   0 OOODD      Ca    2 104    1   23   -1   -1   -1 .. .. 6 1e43   0 0.1 1.7   ....b    |     Z   3.2.1.1;         HYD              ROLASE
166 1axn_   187    45   4   2   0  OOOE       Ca    3    2   40   -1   -1   -1   -1 ....   6 1axn   0 0.2 1.8   ...b     |    ZZ                          CALCIUM/PHOSPHOL
167 1gen_   476   142   4   3   0 OOOO       Ca   45   48   49   -1   -1   -1   -1 ....   7 1gen   0 0.3 2.2   ....     |    ZZZ 3.4.24.24;       HYD              ROLASE
(METAL
168 1qla_A 371    24   5   1   0 OOOOO      Ca    1    1   20    2   -1   -1   -1 ..... 6 1qla   0 0.2 2.2   .....    |     Z   1.3.99.1;        OXI       DOREDUCTASE
169 1nzy_A   49   161   6   0   0 OOOOTQ    Ca 153    1    2    2    3   -1   -1 ......6 1nzy   0 0.4 1.8   ......  |         3.8.1.6;         LYA          SE
170 1fo4_A 867    41   6   0   0 OOOSSO    Ca    3    1    3   33    1   -1   -1 ......6 1fo4   0 0.4 2.1   ......  |         1.1.1.204        OXI        DOREDUCTASE
171 1scj_A 169     5   4   1   0  OOOT       Ca    2    3    0   -1   -1   -1   -1 ....   5 1scj   0 0.4 2.0   ....     |    Z    3.4.21.62;       HYD    ROLASE
172 1hdf_A   19    70   4   2   0  OOSD       Ca   27    2   41   -1   -1   -1   -1 ....   6 1hdf   0 0.2 2.3   ....     |    ZZ                          STRUCTURAL PROTE
173 1nps_A    7    69   4   1   0  OOSN       Ca   30    2   37   -1   -1   -1   -1 ....   5 1nps   0 0.6 1.8   ....     |    Z                            SIGNALING PROTEI
174 1kap_P 253    37   6   0   0 OOTDOD    Ca    2    2   28    2    3   -1   -1 ......6 1kap   1 0.1 1.6   ...b..  |         3.4.24.-;        ZIN        C METALLOPROT
175 1fae_A 185   220   4   2   0 OQED       Ca    0    5 215   -1   -1   -1   -1 ....   6 1fae   0 0.2 2.0   ..b.     |    ZZ   3.2.1.4;         HYD          ROLASE
176 1edm_B   53     0   2   5   0  OS          Ca    0   -1   -1   -1   -1   -1   -1 ..     7 1edm   0 0.2 1.5   ..        | ZZZZZ                          COAGULATION FACT
177 1aru_   185    24   7   0   0  OSDOTOD   Ca    0   17    2    0    3    2   -1 ......7 1aru   0 0.2 1.6   ..b.... |         1.11.1.7;        PER        OXIDASE (DONO
178 1iod_A   41    87   5   1   0  OSEEE      Ca    0    2    4   81   -1   -1   -1 ..... 6 1iod   0 0.3 2.3   ...bb    |     Z   3.4.21.6         HYD       ROLASE/HYDROL
179 1iod_B 241    79   5   2   0  OSQEE      Ca    0    2    4   73   -1   -1   -1 ..... 7 1iod   1 0.3 2.3   ....b    |     ZZ 3.4.21.6         HYD          ROLASE/HYDROL
180 1dl2_A 525     0   2   4   0  OT          Ca    0   -1   -1   -1   -1   -1   -1 ..     6 1dl2   1 0.1 1.5   ..        | ZZZZ   3.2.1.113;       HYD      ROLASE
181 1nbc_B   44    82   6   1   0  OTDOND    Ca    0    2   76    3    1   -1   -1 ......7 1nbc   0 0.2 1.8   ..b...  |      Z                         CEL   LULOSE DEGRAD
182 1pa2_A 170    59   7   0   0 OTDOTOD   Ca    0   51    3    0    3    2   -1 ......7 1pa2   0 0.1 1.5   ....... |         1.11.1.7;        OXI          DOREDUCTASE
183 1brw_A   88   167   5   0   0 OTOOE      Ca    2 153    3    9   -1   -1   -1 ..... 5 1brw   0 0.3 2.1   .....    |         2.4.2.2          TRA       NSFERASE
184 1cse_E    2    79   6   0   0  QDONOO    Ca   39   34    2    2    2   -1   -1 ......6 1cse   0 0.1 1.2   .b....  |         3.4.21.62)       COM          PLEX(SERINE P
185 1gci_      2    79   6   0   0  QDONOO    Ca   39   34    2    2    2   -1   -1 ......6 1gci   0 0.1 0.8   .b....  |         3.4.21.62;       SER      INE PROTEASE
186 1scj_A    2    79   6   0   0  QDONOO    Ca   39   34    2    2    2   -1   -1 ......6 1scj   0 0.2 2.0   .b....  |         3.4.21.62;       HYD         ROLASE
187 2sic_E    2    79   6   0   0  QDONOO    Ca   39   34    2    2    2   -1   -1 ......6 2sic   0 0.2 1.8   .b....  |         3.4.21.14)       COM         PLEX (PROTEIN
188 1tn3_   143    22   4   1   0  QEOD       Ca    7   15    0   -1   -1   -1   -1 ....   5 1tn3   2 0.2 2.0   ....     |    Z                            LECTIN
189 1tf4_A 210    51   5   0   0 SODEO      Ca    1    3    1   46   -1   -1   -1 ..... 5 1tf4   0 0.2 1.9   ..bb.    |         3.2.1.4;         GLY      COSYL HYDROLA
190 1clc_   356    45   5   2   0  TODDO      Ca    2    3    1   39   -1   -1   -1 ..... 7 1clc   0 0.2 1.9   ..b..    |     ZZ 3.2.1.4;         GLY       COSYL HYDROLA
```

```
no   cngpname      nsp  np  nw  nn  dons      met  sd1  sd2   sd3  sd4  sd5  sd6  sd7  metal     startaa    |  conf_efimov
   1  1scf_D   54      4   2   2   0  DD          Ca    4   -1   -1   -1   -1   -1   -1  CA  1022  ASPD   54  aaaaaaaaa
   2  1rdr_   233     96   2   5   0  DD          Ca   96   -1   -1   -1   -1   -1   -1  CA   501  ASP   233  b.bbb-.kbbb
   3  1fsu_    53   248   4   0   3  DD    DN   Ca    1   -1   -1   -1  246    1   -1  CA   604  ASP    53  bbbgbb
   4  1bn8_A  184    43   3   1   0  DDD         Ca   39    4   -1   -1   -1   -1   -1  CA   400  ASPA  184  abbab-kb.babbbb
   5  1alv_B  135    91   4   2   0  DDDN        Ca   88    2    1   -1   -1   -1   -1  CA     8  ASPB  135  b.kjb-bbbaaaaa
   6  2scp_A   16    11   5   1   0  DDDOD       Ca    2    2    2    5   -1   -1   -1  CA   190  ASPA   16  akdakgkgbbbaaaaa
   7  1cdl_A   20    11   5   0   0  DDDOE       Ca    2    2    2    5   -1   -1   -1  CA     1  ASPA   20  aadakgkgbbbaaaaa
   8  1cdl_B  129    11   5   0   0  DDDOE       Ca    2    2    2    5   -1   -1   -1  CA     4  ASPB  129  akbakgkgbbbaaaaa
   9  1acc_   177    11   5   1   0  DDDOE       Ca    2    2    2    5   -1   -1   -1  CA   800  ASP   177  .bdakgkgbbaaaaaa
  10  1g4y_R   20    11   5   1   0  DDDOE       Ca    2    2    2    5   -1   -1   -1  CA  1001  ASPR   20  kabakgkgbbbaaaaa
  11  1sra_   257    11   5   1   0  DDDOE       Ca    2    2    2    5   -1   -1   -1  CA   302  ASP   257  akdakgkgbbbaaaaa
  12  1vrk_A   20    11   5   1   0  DDDOE       Ca    2    2    2    5   -1   -1   -1  CA   151  ASPA   20  aadakgkgbbbaaaaa
  13  1vrk_A  129    11   5   1   0  DDDOE       Ca    2    2    2    5   -1   -1   -1  CA   154  ASPA  129  akbakgkgbbbaaaaa
  14  2pvb_A   90    11   5   1   0  DDDOE       Ca    2    2    2    5   -1   -1   -1  CA   111  ASPA   90  aadkkgkgbbbaaaaa
  15  1acc_   179    56   6   0   0  DDEOOD      Ca    2    7   34    3   10   -1   -1  CA   801  ASP   179  dakgkgb-aaaaa-bbbkkbab-bbaaa
  16  2por_    93     8   4   2   0  DDND        Ca    2    5    1   -1   -1   -1   -1  CA   303  ASP    93  akgdkak.jkbbd
  17  1cdl_B   56    11   6   0   0  DDNNOE      Ca    2    2    0    2    5   -1   -1  CA     2  ASPB   56  aabakgagbbbaaaaa
  18  1cdl_C   56    11   6   0   0  DDNODE      Ca    2    2    2    2    3   -1   -1  CA     2  ASPC   56  aabakgkgbbbaaaaa
  19  1cdl_A   93    11   5   0   0  DDNOE       Ca    2    2    2    5   -1   -1   -1  CA     3  ASPA   93  aadakgkgbbbaaaaa
  20  1cdl_A   56    11   5   0   0  DDNOE       Ca    2    2    2    5   -1   -1   -1  CA     2  ASPA   56  akbakgkgbbbakaaa
  21  1g4y_R   56    11   5   0   0  DDNOE       Ca    2    2    2    5   -1   -1   -1  CA  1002  ASPR   56  adbakgkgbbbaaaka
  22  1rec_   110    11   5   1   0  DDNOE       Ca    2    2    2    5   -1   -1   -1  CA   501  ASP   110  aadakgkgbbbaaaaa
  23  1vrk_A   56    11   5   1   0  DDNOE       Ca    2    2    2    5   -1   -1   -1  CA   152  ASPA   56  akdakgkgbbbaaaaa
  24  1fzc_B  381     4   3   2   0  DDO         Ca    2    2   -1   -1   -1   -1   -1  CA     2  ASPB  381  kbbkbkbbb
  25  2sns_    21    20   3   0   1  DDO         Ca   19    1   -1   -1   -1   -1   -1  CA     1  ASP    21  .akbbbbdbg..b.bbbbbg.bbbb
  26  1kit_   621    62   3   1   0  DDO         Ca   61    1   -1   -1   -1   -1   -1  CA   803  ASP   621  bbbbb-bbbbbb
  27  2mas_B   10   232   4   1   2  DDOD        Ca    5  111  116   -1   -1   -1   -1  CA     6  ASPB   10  bbdbbaaaaa-bbkbb-baaaa
  28  1fzc_C  318     6   4   0   0  DDOO        Ca    2    2    2   -1   -1   -1   -1  CA     1  ASPC  318  k.bdbkba.ba
  29  3fib_   318     6   4   2   0  DDOO        Ca    2    2    2   -1   -1   -1   -1  CA   400  ASP   318  kbbkbkba.ba
  30  1qge_E  241    54   4   2   0  DDOO        Ca   46    4    4   -1   -1   -1   -1  CA   320  ASPE  241  akbak-aaaadgbbggbbb
  31  1ava_A  127    21   5   1   0  DDOOD       Ca   15    1    3    2   -1   -1   -1  CA   502  ASPA  127  akbbkabkgkbbbbk.bbbakbbbbb
```

```
32 1alv_B 180    11   5   1   0 DDSOE     Ca    2    2    2    5   -1   -1   -1 CA     7 ASPB 180 akdkkgkgbbbaaaaa
33 2pvb_A  51    11   6   0   0 DDSOEE    Ca    2    2    2    2    3   -1   -1 CA   110 ASPA  51 aadakgkgbbbaaaaa
34 1alv_A 150    11   5   1   0 DDTOE     Ca    2    2    2    5   -1   -1   -1 CA     2 ASPA 150 akdakgkgbbbaaaaa
35 1f7l_A   8    50   2   3   0 DE         Ca   50   -1   -1   -1   -1   -1   -1 CA   130 ASPA   8 bbbbb-aaaaa
36 2msb_A 161    33   5   1   0 DEDOD     Ca    4   23    5    1   -1   -1   -1 CA     1 ASPA 161 bbdaabbgb-bkgb.k.bkbb
37 1hyo_A 126   107   4   0   2 DEED      Ca   73    2   32   -1   -1   -1   -1 CA 1006 ASPA 126 bbbbd-bbbbbbb-bbbbb
38 1ezm_  136    49   5   1   0 DEEDO     Ca   36    3    8    2   -1   -1   -1 CA   400 ASP  136 baaaa-aaaaaaaa-bbkgbbg
39 1tn3_  116    35   5   1   0 DEOON     Ca    4   27    3    1   -1   -1   -1 CA   182 ASP  116 bbkaabbgb-kk.kkbkbb
40 2sas_   19    11   5   0   0 DNDOD     Ca    2    2    2    5   -1   -1   -1 CA   186 ASP   19 k.dakgkgbbbaaaaa
41 1a2x_A 139    11   5   0   0 DNDOE     Ca    2    2    2    5   -1   -1   -1 CA   161 ASPA 139 aabakgkgbbbaaaaa
42 1a2x_A 103    11   5   0   0 DNDOE     Ca    2    2    2    5   -1   -1   -1 CA   160 ASPA 103 aabakgkgbbbaaaaa
43 2scp_B 104    11   5   0   0 DNDOE     Ca    2    2    2    5   -1   -1   -1 CA   194 ASPB 104 aadak.kgbbbaaaaa
44 2sas_   70    11   5   1   0 DNDOE     Ca    2    2    2    5   -1   -1   -1 CA   187 ASP   70 aabakgkgbbbaaaaa
45 2scp_B 138    11   5   1   0 DNDOE     Ca    2    2    2    5   -1   -1   -1 CA   195 ASPB 138 aadkk.kgbbbaaaaa
46 1gca_  134    71   6   0   0 DNDOQE    Ca    2    2    2    2   63   -1   -1 CA     1 ASP  134 kkdakgkgbbbbb-kbabb
47 1i8a_A  81    13   5   1   0 DNEDD     Ca    2    8    2    1   -1   -1   -1 CA   190 ASPA  81 bbbakgkgbkbbbakbbb
48 1sac_C  58    80   5   0   0 DNEOD     Ca    1   77    1    1   -1   -1   -1 CA     1 ASPC  58 gbbgab-gbbbkbb
49 1sac_A  58   242   6   0   0 DNEODO    Ca    1   77    1    1  162   -1   -1 CA     1 ASPA  58 gbbgab-gbbbkbb-.....
50 1bfd_  428    29   3   1   2 DNO       Ca   27    2   -1   -1   -1   -1   -1 CA   529 ASP  428 bbaaaaaadakaaaaaaakgbdbbbbbbbkgbda
51 1trk_A 157    32   3   1   2 DNO       Ca   30    2   -1   -1   -1   -1   -1 CA   681 ASPA 157 bbaaa-bdkgbba
52 1b0p_B 983    78   5   0   0 DNOOO     Ca    2   71    3    2   -1   -1   -1 CA 1238 ASPB 983 gbbbbbb-aaaakbd.bb
53 1qq9_A   3   263   4   2   0 DODD      Ca    1  258    4   -1   -1   -1   -1 CA   905 ASPA   3 .bbbba-aaaaaaaaa
54 1e43_A 159    45   5   1   0 DODDD     Ca   22    2   19    2   -1   -1   -1 CA   502 ASPA 159 kba.b-gb.bbbb-bbbkaba
55 1oac_A 533   146   5   1   0 DODDO     Ca    1    1  143    1   -1   -1   -1 CA   802 ASPA 533 bbbbbbg-gbbbbb
56 1i8a_A  60    95   5   1   0 DODDO     Ca    2   12   80    1   -1   -1   -1 CA   192 ASPA  60 bb.kbbb-kbb.b-bbbbbk
57 1qho_A  76    26   5   1   0 DODEE     Ca    1    2   22    1   -1   -1   -1 CA   696 ASPA  76 kbbbbkbbdjbgbbdaagkdbkbkkbbbaaa
58 1sra_  222    12   5   1   0 DODOE     Ca    3    2    2    5   -1   -1   -1 CA   301 ASP  222 aabkdbb.gbbbaakka
59 1cvr_A  78   180   3   3   0 DOE       Ca  171    9   -1   -1   -1   -1   -1 CA   501 ASPA  78 b.baa-kbkkb-aaaaa
60 1dx5_J 423    20   6   1   0 DOENOO    Ca    1    2   13    1    3   -1   -1 CA 1001 ASPJ 423 bbbbbaaaajgaba.bbbbbbg.bb
61 1nls_   10     9   4   2   0 DOND      Ca    2    2    5   -1   -1   -1   -1 CA   240 ASP   10 bbbkbbbakkgbbd
62 1qho_A  21    29   6   1   0 DONNOD    Ca    2    3    1   21    2   -1   -1 CA   698 ASPA  21 abbkbkkkadbaadbgkbbkkgkbkkbbkgbbaa
63 1ee6_A  80    23   3   3   0 DOO       Ca    1   22   -1   -1   -1   -1   -1 CA   300 ASPA  80 bk.b.ababbbabb.bbbbbjj.bbgba
64 1d2s_A  50   110   3   3   0 DOO       Ca    2  108   -1   -1   -1   -1   -1 CA   401 ASPA  50 .kbkb.b-babbb
```

```
65 1svy_   222     7   3   3   0 DOO        Ca    5    2   -1   -1   -1   -1   -1 CA     1 ASP   222 aaaaakggabbb
66 1dyk_A2808    68   4   1   0 DOOD       Ca   17   49    2   -1   -1   -1   -1 CA 4001 ASPA2808 gab.b-.bkgk-bkbbbbb
67 1dyk_A2982    73   4   0   0 DOOD       Ca   17   54    2   -1   -1   -1   -1 CA 4002 ASPA2982 a...b-bbkgk-bkbbdb.
68 1fjs_A   70    10   5   1   0 DOOEE      Ca    2    3    2    3   -1   -1   -1 CA   507 ASPA   70 bgbkbad.bagkbbb
69 1gcy_A    1    16   6   0   0 DOOHDE     Ca    1   11    0    3    1   -1   -1 CA   452 ASPA    1    .bb.bbkkgbdkgggkbbb
70 1ga6_A 328    20   5   1   0 DOOOD      Ca    1   15    2    2   -1   -1   -1 CA   374 ASPA 328 bbbbabjbd.bggbjbbb.bjbbaa
71 1i76_A 154    26   6   0   0 DOOODE    Ca    1    2    2   18    3   -1   -1 CA   997 ASPA 154 bbkjbjjba.bbbbb.b.abgbbbbbkkbbb
72 1hdf_B   45    47   4   2   0 DOOS       Ca   17   28    2   -1   -1   -1   -1 CA 1101 ASPB   45 abgkb-bbabb-bg.bk..
73 1edm_C   47    18   5   0   0 DOQDO      Ca    1    2   14    1   -1   -1   -1 CA     1 ASPC   47 ..bbbkkakgkbbggbbbbbbbg
74 1kap_P 446      8   5   1   0 DSDOD      Ca    2    2    2    2   -1   -1   -1 CA   621 ASPP 446 bbbaagkgbbabb
75 2sas_   115    11   5   0   0 DSDOE      Ca    2    2    2    5   -1   -1   -1 CA   188 ASP   115 kadakgkgbbbaaaaa
76 1c7k_A   76      2   2   4   0 DT          Ca    2   -1   -1   -1   -1   -1   -1 CA   134 ASPA   76 abbaaaa
77 1hei_A 441      4   2   1   0 DT          Ca    4   -1   -1   -1   -1   -1   -1 CA   633 ASPA 441 ag.a.bbbb
78 2cbl_A 229    11   5   1   0 DTNOE      Ca    2    2    2    5   -1   -1   -1 CA   352 ASPA 229 aadakgkgbbbaaaaa
79 1d0b_A   55      4   2   3   0 ED          Ca    4   -1   -1   -1   -1   -1   -1 CA   202 GLUA   55 aaaaaaaak
80 2por_    80    28   2   4   0 ED          Ca   28   -1   -1   -1   -1   -1   -1 CA   302 GLU    80 aaaaakjbbbbbjakgdkak.jkbbdbabjgga
81 2msb_A 165    29   2   4   0 ED          Ca   29   -1   -1   -1   -1   -1   -1 CA     3 GLUA 165 abbgbbbbakj.bbabbdbbbgbbkgb.k.bkbb
82 1g0h_A   65   136   4   0   2 EDDD       Ca   16    3 117   -1   -1   -1   -1 CA   290 GLUA   65 bbad.-bbbabbga-aaaka
83 1cb8_A 405    12   4   2   0 EDDO       Ca    2    9    1   -1   -1   -1   -1 CA 3000 GLUA 405 kbbkkdbbgbb.bbbab
84 1g0h_B 365    18   3   0   1 EDO        Ca   16    2   -1   -1   -1   -1   -1 CA   590 GLUB 365 bbad..bbbkgbkbbbbbbabbg
85 1atl_A    9   191   4   1   0 EDON       Ca   84 104    3   -1   -1   -1   -1 CA   403 GLUA    9 bbbbb-b.abb-baakabb.
86 1sac_C 136    12   3   0   0 EDQ        Ca    2   10   -1   -1   -1   -1   -1 CA     2 GLUC 136 gbbbkbbg.gbbakkbb
87 1blx_A   69      3   2   4   0 EE          Ca    3   -1   -1   -1   -1   -1   -1 CA     0 GLUA   69 aaakagba
88 1i7o_A 276    31   3   3   0 EED        Ca    2   29   -1   -1   -1   -1   -1 CA 1001 GLUA 276 bbbbbbb-bbdbb
89 1cvr_A 293      9   3   1   0 EEH        Ca    3    6   -1   -1   -1   -1   -1 CA   686 GLUA 293 aaaaaaaakgkbak
90 1cvr_A 161   131   3   3   0 EHD        Ca    5 126   -1   -1   -1   -1   -1 CA   648 GLUA 161 ggbbaaaaaa-aaaaa
91 1g72_A 171    84   2   0   3 EN         Ca   84   -1   -1   -1   -1   -1   -1 CA   702 GLUA 171 .akkg-bbbbb
92 1byf_B   86    22   5   0   2 ENDOD      Ca    3   18    1    0   -1   -1   -1 CA   201 GLUB   86 bgbbkbbakbbbbbbbbaakgbbbbbb
93 2msb_A 185    21   6   0   0 ENENOD    Ca    2    6   12    1    0   -1   -1 CA     2 GLUA 185 bgbbkgb.k.bkbbbbbakgbbbbbb
94 1cru_B 253    10   2   4   0 EO         Ca   10   -1   -1   -1   -1   -1   -1 CA   901 GLUB 253 gbbbbbbbbgbbkgb
95 1bag_   276    37   2   3   0 EO          Ca   37   -1   -1   -1   -1   -1   -1 CA     8 GLU   276 kagak-jggba
96 1oac_A 573    99   4   2   0 EODE       Ca   94    3    2   -1   -1   -1   -1 CA   803 GLUA 573 kbaaa-aaakakgbbk
97 1g5c_F   97    24   3   3   0 EOO        Ca   21    3   -1   -1   -1   -1   -1 CA 1009 GLUF   97 bbaaaaaaaaaakgbbaaaaaaaaaabbb
```

```
 98 1ava_B 108    9  4  1  0 EOOD     Ca   3   2   4  -1  -1  -1  -1 CA  501 GLUB 108 bbbkgbkbkbgdbb
 99 1sbw_A  70   10  4  2  0 EOOE     Ca   2   3   5  -1  -1  -1  -1 CA  801 GLUA  70 ..b.bkkbbk.gbbb
100 2btc_E  70   10  4  2  0 EOOE     Ca   2   3   5  -1  -1  -1  -1 CA  700 GLUE  70 ..babkkbbkgkbbb
101 1avw_A  70   10  5  1  0 EOOEE    Ca   2   3   2   3  -1  -1  -1 CA  700 GLUA  70 ..babkkbbkgkbbb
102 1fs7_A 217   59  4  2  0 EOOQ     Ca   1  56   2  -1  -1  -1  -1 CA  651 GLUA 217 dbbbbb-bbbbb.d
103 1pyt_C 470   10  5  0  0 EOOQE    Ca   2   3   2   3  -1  -1  -1 CA  650 GLUC 470 bgb.bkkbbkgabbb
104 1gcy_A 116   81  5  1  0 NDODO    Ca  35   3   8  35  -1  -1  -1 CA  451 ASNA 116 baabb-kk.kbbbb-kkbbb-kkkbb
105 1ava_A  91   92  5  1  0 NDODO    Ca  47   3   7  35  -1  -1  -1 CA  500 ASNA  91 bkabb-bbbk.bbb-bbbbb-kkkbb
106 2por_  116   24  4  1  0 NDOO     Ca  20   2   2  -1  -1  -1  -1 CA  304 ASN  116 bgdbbbbbbbbb.kbbbbbb.bkgbb.db
107 1jg8_D 330    3  2  1  0 NE       Ca   3  -1  -1  -1  -1  -1  -1 CA  906 ASND 330 baaaaaaa
108 1e43_A 444    3  2  5  0 NE       Ca   3  -1  -1  -1  -1  -1  -1 CA  503 ASNA 444 akkbgbbb
109 1i9z_A 568   29  2  4  0 NE       Ca  29  -1  -1  -1  -1  -1  -1 CA  301 ASNA 568 bkggbbbbkdkaaaadabdbbbbabbbbbbbgab
110 1djx_B 312   78  4  1  1 NEDE     Ca  29   2  47  -1  -1  -1  -1 CA    2 ASNB 312 bbgak-bbbbbbb-bbbbk
111 1e29_A  49    1  2  6  0 NN       Ca   1  -1  -1  -1  -1  -1  -1 CA  225 ASNA  49 kkbbgd
112 1egi_A 747    1  3  1  0 NOD      Ca   1   0  -1  -1  -1  -1  -1 CA  801 ASNA 747 bbbbbb
113 1e43_A 102  133  5  1  0 NODDO    Ca  92   0   6  35  -1  -1  -1 CA  501 ASNA 102 bkabb-bkabab.bbbb-akkbb
114 1bag_  101   79  4  1  0 NODO     Ca  36   9  34  -1  -1  -1  -1 CA    6 ASN  101 bkabb-aaakb-bbbbd-kkkbb
115 7taa_  121   89  4  3  0 NODO     Ca  41  13  35  -1  -1  -1  -1 CA  477 ASN  121 bka.b-aaakb-bbbbb-akkbb
116 1clv_A  98   91  4  3  0 NODO     Ca  48   9  34  -1  -1  -1  -1 CA  601 ASNA  98 bkab.-aaakb-bbbbb-aakbb
117 1qho_A 131  101  4  3  0 NODO     Ca  53  14  34  -1  -1  -1  -1 CA  697 ASNA 131 dkab.-aaakb-bbbbb-kkkbb
118 1nps_A  36   43  3  3  0 NOS      Ca  17  26  -1  -1  -1  -1  -1 CA  145 ASNA  36 bbgab-bbabb-.bkbb
119 1e29_A  35    0  2  4  0 OD       Ca   0  -1  -1  -1  -1  -1  -1 CA  226 ASPA  35 aaaaa
120 1bag_  169    2  2  3  0 OD       Ca   2  -1  -1  -1  -1  -1  -1 CA    7 GLY  169 ak.bk.b
121 1d0b_A  49    2  2  5  0 OD       Ca   2  -1  -1  -1  -1  -1  -1 CA  201 PROA  49 ababaaa
122 1qh4_D  41    3  2  0  0 OD       Ca   3  -1  -1  -1  -1  -1  -1 CA  382 LYSD  41 aaakkkbb
123 3std_C  51    4  2  2  0 OD       Ca   4  -1  -1  -1  -1  -1  -1 CA  502 ARGC  51 bdaaakgbb
124 1wdc_C  19    8  6  1  0 ODDODO   Ca   0   3   1   2   2  -1  -1 CA  501 ASPC  19 aaaakk.b.gbbb
125 1djx_A 651   26  3  2  0 ODN      Ca   2  24  -1  -1  -1  -1  -1 CA    3 ILEA 651 abbbdbbbbbbbb.baakb.bbbbbbbbgbk
126 1clc_  520    5  3  3  0 ODO      Ca   3   2  -1  -1  -1  -1  -1 CA  593 SER  520 aaakkbkbab
127 1e54_A  29  280  3  0  0 ODO      Ca  35 245  -1  -1  -1  -1  -1 CA  500 LEUA  29 b.bbb-baakg-bbga.
128 1ajj_   22   14  6  0  0 ODODDE   Ca   3   2   2   6   1  -1  -1 CA   73 TRP   22 akkbgkgbbabakgakkbg
129 1i8a_A  10  120  5  1  0 ODODE    Ca   2   2   2 114  -1  -1  -1 CA  191 VALA  10 bbbb.gbbbaa-bbbbb
130 1aru_   57   22  5  2  0 ODODS    Ca   0  18   2   2  -1  -1  -1 CA  346 ASP   57 aaaakdbbaaaaakgbbk.kjdkgaaa
```

```
131 1kap_P 288   41  4  2  0 ODOE     Ca   2  37   2  -1  -1  -1  -1 CA  615 GLYP 288 k.jbbbb-gbbbabb
132 1cvr_A 100    7  4  2  0 ODOE     Ca   3   2   2  -1  -1  -1  -1 CA  477 VALA 100 babggbkbbabb
133 1alv_B 107   10  4  2  0 ODOE     Ca   3   2   5  -1  -1  -1  -1 CA    5 ALAB 107 aaajakgbbbaaaaa
134 1tf4_B 504   71  5  1  0 ODOND    Ca   2  65   3   1  -1  -1  -1 CA 3004 THRB 504 bbbbbbg-gbbkakbak
135 1pa2_A  43    9  6  1  0 ODOODS   Ca   0   3   2   2   2  -1  -1 CA  307 ASPA  43 aaaakbgjdkgaak
136 1e8u_A 261   35  5  0  0 ODOOO    Ca   0   3   2  30  -1  -1  -1 CA 1002 ASPA 261 aaaakbabbb-kbkbb
137 4sgb_E 120  123  5  1  0 ODOOY    Ca   0   1 122   0  -1  -1  -1 CA    8 GLYE 120 ..bbab-bb...
138 1e8u_B 261   35  6  0  0 ODOSOO   Ca   0   3   0   2  30  -1  -1 CA 1003 ASPB 261 aaakkbabbb-kbabb
139 1g5c_C 147    0  2  3  0 OE       Ca   0  -1  -1  -1  -1  -1  -1 CA 1001 GLUC 147 bbkkb
140 1ayo_B 120    1  2  3  0 OE       Ca   1  -1  -1  -1  -1  -1  -1 CA    1 ASPB 120 bbgdbb
141 1c1y_B 123    2  2  2  0 OE       Ca   2  -1  -1  -1  -1  -1  -1 CA  173 GLYB 123 kbgbbbb
142 1qhd_A 151  189  2  2  0 OE       Ca 189  -1  -1  -1  -1  -1  -1 CA  603 THRA 151 bbbbb-kabka
143 1dyo_A  14  135  5  0  0 OEOOD    Ca   2  23   3 107  -1  -1  -1 CA  300 THRA  14 bbagdg.-.abjkbbb-bbbgb
144 1fmj_A 118    4  3  0  0 OND      Ca   3   1  -1  -1  -1  -1  -1 CA  402 GLUA 118 kaakakkkb
145 1clc_  236   10  5  1  0 ONODD    Ca   3   2   2   3  -1  -1  -1 CA  591 GLU  236 bakkkbkbbaaaaaa
146 1cru_B 247    1  2  2  3 OO       Ca   1  -1  -1  -1  -1  -1  -1 CA  908 GLYB 247 bbb.kb
147 1c8n_A 173    2  2  0  0 OO       Ca   2  -1  -1  -1  -1  -1  -1 CA  305 SERA 173 aaakbbg
148 1kap_P 370   30  4  2  0 OODD     Ca   2   2  26  -1  -1  -1  -1 CA  617 GLYP 370 bajbjbbbbbbkbaaadbaadbbbbkdbbbgabbb
149 1lpb_B 187    8  4  2  0 OODD     Ca   3   2   3  -1  -1  -1  -1 CA  450 GLUB 187 bbaakbb.akkba
150 1cru_B 269   40  4  2  0 OODE     Ca   2   2  36  -1  -1  -1  -1 CA  902 ALAB 269 abbjbbkgb-bbakk
151 1f8e_A 293   54  4  2  0 OODO     Ca   4  27  23  -1  -1  -1  -1 CA  999 ASPA 293 bbbaa.bkb-bkbbb-bg..b
152 1kap_P 361   29  5  1  0 OODOD    Ca   2   2  18   7  -1  -1  -1 CA  619 GLYP 361 bajbjbbbbbajbjbbbbbbkbaaadbaadbbbb
153 1kap_P 334   22  6  0  0 OODOOD   Ca   2   2  13   2   3  -1  -1 CA  616 GLYP 334 bb.kbbbbbbbbkbbbbbbajbjbbbb
154 1kap_P 352   22  6  0  0 OODOOD   Ca   2   2  13   2   3  -1  -1 CA  620 GLYP 352 bajbjbbbbbajbjbbbbbajbjbbbb
155 1sra_  241    5  3  2  0 OOE      Ca   2   3  -1  -1  -1  -1  -1 CA  303 PRO  241 kbakbbgkkk
156 1ej8_A 124   44  3  3  0 OON      Ca  44   0  -1  -1  -1  -1  -1 CA  218 PROA 124 gbabb-akgbb
157 1kit_  253   60  6  0  0 OONDOT   Ca   3   0  33  24   0  -1  -1 CA  802 ALA  253 bgbkd.bb-kbbbb-bbkkb
158 1kap_P 343   22  6  0  0 OONOOD   Ca   2   2  13   2   3  -1  -1 CA  618 ASNP 343 bbbkbbbbbbajbjbbbbbajbjbbbb
159 1dmu_A  99    5  3  3  0 OOO      Ca   2   3  -1  -1  -1  -1  -1 CA  125 SERA  99 aaakbakbbb
160 1a0t_P 454    8  3  0  0 OOO      Ca   3   5  -1  -1  -1  -1  -1 CA   10 ASNP 454 akakabakbabgb
161 1poc_    8   27  4  0  2 OOOD     Ca   2   2  23  -1  -1  -1  -1 CA  501 TRP    8 bkkbjbjbdbkbakbbbadaaaaaaaaaaakb
162 1axn_   32   44  4  0  0 OOOD     Ca   2   2  40  -1  -1  -1  -1 CA  351 ILE   32 aaabjbjbb-aaabb
163 2sic_E 169   28  4  2  0 OOOD     Ca   5  21   2  -1  -1  -1  -1 CA  502 GLYE 169 bkaabakbbbbbbbbkkgbbbaabbb.kkbabb
```

```
164 1i76_A 137    36   4   2   0 OOOD       Ca   32    2    2   -1   -1   -1   -1 CA   996 ASPA 137 bbabb-.b.abgbbb
165 1e43_A 300   130   5   1   0 OOODD      Ca    2 104    1   23   -1   -1   -1 CA   504 GLYA 300 kggkbbk-kbkabb-bbk.b
166 1axn_   187    45   4   2   0 OOOE       Ca    3    2   40   -1   -1   -1   -1 CA   353 GLY   187 ak.akbbjbb-aaabb
167 1gen_   476   142   4   3   0 OOOO       Ca   45   48   49   -1   -1   -1   -1 CA   302 ASP   476 bbabb-bba.b-.babb-gba.b
168 1qla_A 371    24   5   1   0 OOOOO      Ca    1    1   20    2   -1   -1   -1 CA     9 SERA 371 .bbab.bbbbkkgbbabbgbbbbjakbda
169 1nzy_A  49   161   6   0   0 OOOOTQ    Ca 153    1    2    2    3   -1   -1 CA   271 GLYA  49 kbk.b-aaaaabbaaaaaa
170 1fo4_A 867    41   6   0   0 OOOSSO    Ca    3    1    3   33    1   -1   -1 CA 4009 ALAA 867 bbajk.a.kaaa-bbbb.b
171 1scj_A 169     5   4   1   0 OOOT       Ca    2    3    0   -1   -1   -1   -1 CA   382 ALAA 169 bkaabaabbb
172 1hdf_A  19    70   4   2   0 OOSD       Ca   27    2   41   -1   -1   -1   -1 CA 1101 LYSA  19 bbabb-bgkbabb-dbg.b
173 1nps_A    7    69   4   1   0 OOSN       Ca   30    2   37   -1   -1   -1   -1 CA    90 TYRA    7 bbabb-bgabkbb-kbg.b
174 1kap_P 253    37   6   0   0 OOTDOD    Ca    2    2   28    2    3   -1   -1 CA   614 ARGP 253 kbgkbbbbb-bbbk.jbbbb
175 1fae_A 185   220   4   2   0 OQED       Ca    0    5 215   -1   -1   -1   -1 CA 2000 GLNA 185 jakbbakkbb-bkdbb
176 1edm_B  53     0   2   5   0 OS          Ca    0   -1   -1   -1   -1   -1   -1 CA     3 SERB  53 kbgga
177 1aru_   185    24   7   0   0 OSDOTOD   Ca    0   17    2    0    3    2   -1 CA   347 SER   185 akkdbbbakabkadbgbbkbkdkkbbbka
178 1iod_A  41    87   5   1   0 OSEEE      Ca    0    2    4   81   -1   -1   -1 CA   501 SERA  41 bbbbk.aaaaa-bbb..
179 1iod_B 241    79   5   2   0 OSQEE      Ca    0    2    4   73   -1   -1   -1 CA   502 SERB 241 bbbbkbaaaaa-bbbbb
180 1dl2_A 525     0   2   4   0 OT          Ca    0   -1   -1   -1   -1   -1   -1 CA   901 THRA 525 bbkkg
181 1nbc_B  44    82   6   1   0 OTDOND    Ca    0    2   76    3    1   -1   -1 CA     1 THRB  44 b.bag.b-bbbkkdbak
182 1pa2_A 170    59   7   0   0 OTDOTOD   Ca    0   51    3    0    3    2   -1 CA   308 THRA 170 kkkbj-babkabakbbbka
183 1brw_A  88   167   5   0   0 OTOOE      Ca    2 153    3    9   -1   -1   -1 CA 3001 GLYA  88 .bgbkaa-bbbjabb.-aaaaa
184 1cse_E    2    79   6   0   0 QDONOO    Ca   39   34    2    2    2   -1   -1 CA   430 GLNE    2  .bbb-bakbb-kbbb.kbjbbb
185 1gci_     2    79   6   0   0 QDONOO    Ca   39   34    2    2    2   -1   -1 CA   277 GLN     2  .bbb-bkkbb-.bbb.kbjbbb
186 1scj_A    2    79   6   0   0 QDONOO    Ca   39   34    2    2    2   -1   -1 CA   381 GLNA    2  .bbb-bkkbb-dbbb.kb.bbb
187 2sic_E    2    79   6   0   0 QDONOO    Ca   39   34    2    2    2   -1   -1 CA   501 GLNE    2  .bbb-bakbb-dbbb.dbjbbb
188 1tn3_   143    22   4   1   0 QEOD       Ca    7   15    0   -1   -1   -1   -1 CA   183 GLN   143 bbbbkk.kkbkbbbbbaakggbbbbbb
189 1tf4_A 210    51   5   0   0 SODEO      Ca    1    3    1   46   -1   -1   -1 CA 3001 SERA 210 bakjkaaaaa-.kabk
190 1clc_   356    45   5   2   0 TODDO      Ca    2    3    1   39   -1   -1   -1 CA   592 THR   356 gbbabkaaaaa-.aabk
```

```
no  cngpname     nsp  np  nw  nn  dons      met  sd1  sd2  sd3  sd4  sd5  sd6  sd7  his   cn  pdb  cn2  rms  res   carbi    othdon   ecno

  1 8ruc_A 203     1   2   0   3   DE        Mg   -1    1   -1   -1   -1   -1   -1 ...    6  8ruc   0  0.3  1.6   ...      |   UUU   4.1.1.39         LYA   SE  (CARBON-CA
  2 1bpy_A 190     2   2   1   3  DD         Mg    2   -1   -1   -1   -1   -1   -1 ..     6  1bpy   0  0.3  2.2   ..       |  UUUZ   2.7.7.7;         COM      PLEX  (NUCLEOT
  3 1eqo_A  95     2   2   1   3  DD         Mg    2   -1   -1   -1   -1   -1   -1 ..     6  1eqo   0  0.2  1.3   ..       |  UUUZ   2.7.6.3;         TRA   NSFERASE
  4 1uby_  117     4   2   0   1  DD         Mg    4   -1   -1   -1   -1   -1   -1 ..     3  1uby   0  0.4  2.4   ..       |  U       2.5.1.1;         TRANSFERASE
  5 1bl3_B  64    52   2   2   0  DD         Mg   52   -1   -1   -1   -1   -1   -1 ..     4  1bl3   0  0.6  2.0   ..       |  ZZ                         INTEGRATION
  6 1t7p_A 475   179   2   2   1  DD         Mg  179   -1   -1   -1   -1   -1   -1 ..     5  1t7p   1  0.3  2.2   ..       |  UZZ    2.7.7.7;         COM       PLEX(HYDROLAS
  7 1bpy_A 190    66   3   0   1  DDD        Mg    2   64   -1   -1   -1   -1   -1 ...    4  1bpy   0  0.3  2.2   b..      |   U      2.7.7.7;         COM      PLEX  (NUCLEOT
  8 1obw_A  65    37   3   3   0  DDD        Mg    5   32   -1   -1   -1   -1   -1 ...    6  1obw   0  0.2  1.9   ...      |   ZZZ   3.6.1.1;         HYD       ROLASE
  9 1auk_   29   253   4   0   0  DDDN       Mg    1  251    1   -1   -1   -1   -1 ....   4  1auk   0  0.3  2.1   ..b.     |         3.1.6.8;         HYD    ROLASE
 10 1wdc_B  28    11   5   1   0  DDDOD      Mg    2    2    2    5   -1   -1   -1 ..... 6  1wdc   1  0.3  2.0   .....    |      Z                      MUS  CLE  PROTEIN
 11 1vid_  141    29   3   1   2  DDN        Mg   28    1   -1   -1   -1   -1   -1 ...    6  1vid   0  0.2  2.0   ...      |   UUZ   2.1.1.6;         TRA     NSFERASE  (MET
 12 1dqn_A 125     1   2   2   2  DE         Mg    1   -1   -1   -1   -1   -1   -1 ..     6  1dqn   0  0.4  1.8   ..       |  UUZZ   2.4.2.8;         TRA     NSFERASE
 13 1gsa_  273     8   2   1   3  DE         Mg    8   -1   -1   -1   -1   -1   -1 ..     6  1gsa   0  0.3  2.0   ..       |  UUUZ   6.3.2.3;         LIG  ASE
 14 1iow_  257    13   2   1   3  DE         Mg   13   -1   -1   -1   -1   -1   -1 ..     6  1iow   0  0.2  1.9   ..       |  UUUZ   6.3.2.4;         LIG     ASE
 15 1qr0_A 107    44   3   1   2  DEE        Mg    2   42   -1   -1   -1   -1   -1 ...    6  1qr0   0  0.4  1.9   ...      |   JJZ                       TRANSFERASE
 16 1mdl_  195    52   3   1   2  DEE        Mg   26   26   -1   -1   -1   -1   -1 ...    6  1mdl   1  0.2  1.9   ...      |   UUZ   5.1.2.2;         ISO      MERASE
 17 1ec7_A 235    54   3   3   0  DEN        Mg   25   29   -1   -1   -1   -1   -1 ...    6  1ec7   0  0.2  1.9   ...      |   ZZZ   4.2.1.40;        LYA         SE
 18 1bx4_A 130     1   2   0   4  DN         Mg    1   -1   -1   -1   -1   -1   -1 ..     6  1bx4   0  0.2  1.5   ..       |  UUUU   2.7.1.20;        TRA       NSFERASE
 19 1g97_A 102   125   2   2   2  DN         Mg  125   -1   -1   -1   -1   -1   -1 ..     6  1g97   0  0.1  2.0   ..       |  UUZZ   2.7.7.23;        TRA           NSFERASE
 20 1ig5_A  54     6   4   2   0  DNDO       Mg    2    2    2   -1   -1   -1   -1 ....   6  1ig5   0  0.2  1.5   ....     |    ZZ                       MET BINDING  PR
 21 1pox_A 447    29   3   1   2  DNO        Mg   27    2   -1   -1   -1   -1   -1 ...    6  1pox   0  0.2  2.1   ...      |   UUZ   1.2.3.3) M       OXI           DOREDUCTASE
(O
 22 1pvd_A 444    29   3   1   2  DNO        Mg   27    2   -1   -1   -1   -1   -1 ...    6  1pvd   0  0.2  2.3   ...      |   UUZ   4.1.1.1)         LYA         SE  (CARBON-CA
 23 1zpd_A 440    29   3   1   2  DNO        Mg   27    2   -1   -1   -1   -1   -1 ...    6  1zpd   0  0.2  1.9   ...      |   UUZ   4.1.1.1;         ALC        OHOL  FERMENTA
 24 1qs0_A 213    31   3   1   2  DNO        Mg   29    2   -1   -1   -1   -1   -1 ...    6  1qs0   1  0.2  2.4   ...      |   UUZ                       OXI DOREDUCTASE
 25 1qf5_A  13    27   2   0   5  DO         Mg   27   -1   -1   -1   -1   -1   -1 ..     7  1qf5   0  0.4  2.0   ..       |  UUUJJ  6.3.4.4          LIG    ASE
 26 1fiu_B 140    46   2   2   2  DO         Mg   46   -1   -1   -1   -1   -1   -1 ..     6  1fiu   0  0.2  1.6   ..       |  UUZZ   3.1.21.4;        HYD       ROLASE/DNA
 27 1t7p_A 475   179   3   0   3  DOD        Mg    1  178   -1   -1   -1   -1   -1 ...    6  1t7p   0  0.2  2.2   ...      |   UUU   2.7.7.7;         COM           PLEX(HYDROLAS
 28 1hyo_A 233    24   5   0   0  DOOOT      Mg    1   19    3    1   -1   -1   -1 ..... 5  1hyo   0  0.4  1.3   .....    |          3.7.1.2;         HYD       ROLASE
 29 1f1z_A 114    17   3   3   0  DQO        Mg   16    1   -1   -1   -1   -1   -1 ...    6  1f1z   0  0.2  2.4   ...      |   ZZZ                       DNA  BINDING  PROT
 30 1ew2_A  42   269   3   3   0  DSE        Mg  113  156   -1   -1   -1   -1   -1 ...    6  1ew2   0  0.3  1.8   ...      |   ZZZ   3.1.3.1          HYD            ROLASE
```

```
31 1qb7_A 146     8  2  1  0 DT        Mg    8  -1  -1  -1  -1  -1  -1 ..     3 1qb7  1 0.6 1.5  ..       | Z                     TRA NSFERASE
32 5eau_  444     8  3  1  0 DTE       Mg    4   4  -1  -1  -1  -1  -1 ...    4 5eau  0 0.5 2.2  ...      |  Z                    ISO PRENOID SYNTH
33 1b0p_A 963    30  3  1  2 DTO       Mg   28   2  -1  -1  -1  -1  -1 ...    6 1b0p  0 0.3 2.3  ...      |  UUZ  1.2.7.1;        OXI      DOREDUCTASE
34 1g29_2 293    63  3  0  0 DYD       Mg    2  61  -1  -1  -1  -1  -1 ...    3 1g29  0 0.6 1.9  ...      |                       SUG AR BINDING PR
35 1g29_2 292     1  2  0  1 ED        Mg    1  -1  -1  -1  -1  -1  -1 ..     3 1g29  1 0.3 1.9  ..       | J                     SUG AR BINDING PR
36 1a49_B 871    24  2  1  2 ED        Mg   24  -1  -1  -1  -1  -1  -1 ..     5 1a49  0 0.2 2.1  ..       | UUZ   2.7.1.40;       TRA      NSFERASE
37 1dxe_A 153    26  2  2  1 ED        Mg   26  -1  -1  -1  -1  -1  -1 ..     5 1dxe  0 0.2 1.8  ..       | UZZ   4.1.2.20        CLA      SS II ALDOLAS
38 1d8c_A 427    28  2  2  2 ED        Mg   28  -1  -1  -1  -1  -1  -1 ..     6 1d8c  0 0.2 2.0  ..       | UUZZ  4.1.3.2;        LYA       SE
39 1d3y_A 197    52  2  4  0 ED        Mg   52  -1  -1  -1  -1  -1  -1 ..     6 1d3y  0 0.2 2.0  ..       | ZZZZ  5.99.1.3;       ISO      MERASE
40 1g29_1 292    67  3  0  1 EDK       Mg    1  66  -1  -1  -1  -1  -1 ...    4 1g29  0 0.5 1.9  ...      |  J                    SUG AR BINDING PR
41 1qgx_A  72    72  3  1  2 EDO       Mg   70   2  -1  -1  -1  -1  -1 ...    6 1qgx  0 0.3 1.6  ...      |  UUZ  3.1.3.7;        HYD       ROLASE
42 1alo_  899     4  2  3  0 EE        Mg    4  -1  -1  -1  -1  -1  -1 ..     5 1alo  0 0.4 2.0  ..       | ZZZ                   OXI DOREDUCTASE
43 1eyz_A 267    12  2  1  2 EE        Mg   12  -1  -1  -1  -1  -1  -1 ..     5 1eyz  0 0.2 1.8  b.       | UUZ   2.1.2.-;        TRA     NSFERASE
44 1i7q_A 361   137  2  2  2 EE        Mg  137  -1  -1  -1  -1  -1  -1 ..     6 1i7q  0 0.4 2.0  ..       | UUZZ  4.1.3.27;       LYA         SE
45 1dp0_A 416    45  3  3  0 EHE       Mg    2  43  -1  -1  -1  -1  -1 .d.    6 1dp0  0 0.2 1.7  ...      |  ZZZ  3.2.1.23;       HYD         ROLASE
46 1iow_  270     2  2  1  2 EN        Mg    2  -1  -1  -1  -1  -1  -1 ..     5 1iow  0 0.3 1.9  b.       | UUZ   6.3.2.4;        LIG ASE
47 1gsa_  281     2  2  1  3 EN        Mg    2  -1  -1  -1  -1  -1  -1 ..     6 1gsa  1 0.4 2.0  b.       | UXUZ  6.3.2.3;        LIG  ASE
48 2hgs_A 144   224  3  0  2 ENE       Mg    2 222  -1  -1  -1  -1  -1 ...    5 2hgs  0 0.2 2.1  b..      |  UU   6.3.2.3;        AMI        NE/CARBOXYLAT
49 1inp_   79    76  2  0  0 EO        Mg   76  -1  -1  -1  -1  -1  -1 ..     2 1inp  0 0.4 2.3  ..       |       3.1.3.57)       HYD ROLASE(PHOSPH
50 1ew2_A 216    69  4  1  0 EOED      Mg   53   1  15  -1  -1  -1  -1 ....   5 1ew2  0 0.3 1.8  b..b     |   Z   3.1.3.1         HYD         ROLASE
51 1g8f_A  46   120  3  2  0 EOH       Mg  118   2  -1  -1  -1  -1  -1 ..d    5 1g8f  0 0.5 2.0  b..      |  ZZ   2.7.7.4         TRA       NSFERASE
52 1h2r_L  62   490  3  3  0 EOH       Mg  436  54  -1  -1  -1  -1  -1 ..e    6 1h2r  0 0.1 1.4  ...      |  ZZZ  1.12.2.1;       OXI         DOREDUCTASE
53 1a6o_  161    14  2  0  0 ND        Mg   14  -1  -1  -1  -1  -1  -1 ..     2 1a6o  1 0.4 2.1  ..       |       2.7.1.37;       TRANSFERASE
54 1alo_  649    44  3  1  0 OEO       Mg    2  42  -1  -1  -1  -1  -1 ...    4 1alo  0 0.3 2.0  .b.      |  Z                    OXI DOREDUCTASE
55 1e3m_A  99     3  2  0  4 OO        Mg    3  -1  -1  -1  -1  -1  -1 ..     6 1e3m  0 0.3 2.2  ..       | UUUU                    DNA  BINDING
56 1b25_A  92    89  2  2  2 OO        Mg   89  -1  -1  -1  -1  -1  -1 ..     6 1b25  0 0.1 1.9  ..       | UUZZ                    MOC  O
57 1bfd_  117     3  3  0  0 OOO       Mg    1   2  -1  -1  -1  -1  -1 ...    3 1bfd  1 0.3 1.6  ...      |       4.1.1.7;        LYA SE
58 1ryp_M 183     6  3  1  0 OOO       Mg    3   3  -1  -1  -1  -1  -1 . .    4 1ryp  0 0.3 1.9  ...      |  Z    3.4.99.46;      MUL    TICATALYTIC P
59 1ryp_H 163     6  3  0  0 OOO       Mg    3   3  -1  -1  -1  -1  -1 ...    3 1ryp  0 0.5 1.9  ...      |       3.4.99.46;      MUL  TICATALYTIC P
60 1ryp_J 166     6  3  2  0 OOO       Mg    3   3  -1  -1  -1  -1  -1 ...    5 1ryp  0 0.3 1.9  ...      |  ZZ   3.4.99.46;      MUL    TICATALYTIC P
61 1ewk_B  89     7  3  2  0 OOO       Mg    6   1  -1  -1  -1  -1  -1 ...    5 1ewk  1 0.3 2.2  ...      |  ZZ                     SIG NALING PROTEI
62 1ewk_A  89     7  4  2  0 OOOO      Mg    3   3   1  -1  -1  -1  -1 ....   6 1ewk  0 0.4 2.2  ....     |   ZZ                    SIG NALING PROTEI
63 1dp0_D  15   178  5  0  0 OOOQD     Mg    3   3 142  30  -1  -1  -1 ..... 5 1dp0  0 0.3 1.7  ....b    |       3.2.1.23;       HYD          ROLASE
```

```
 64 1ezw_A   56     1   2   1   0 OS         Mg    1   -1   -1   -1   -1   -1   -1 ..     3 1ezw 20 0.5 1.6   ..        | Z                          OXI DOREDUCTASE
 65 3prn_    28    29   2   3   0 SD         Mg   29   -1   -1   -1   -1   -1   -1 ..     5 3prn   0 0.3 1.9   ..        | ZZZ                        MEM BRANE PROTEIN
 66 3pmg_B 116   175   4   1   1 SDDD      Mg 171    2    2   -1   -1   -1   -1 ....  6 3pmg   0 0.2 2.4   ....      |    UZ   5.4.2.2;
PHOSPHOTRANSFERA
 67 1qc5_B 342   101   3   3   0 SSD       Mg    2   99   -1   -1   -1   -1   -1 ...    6 1qc5   0 0.2 2.0   ...      |  ZZZ                        CEL L ADHESION
 68 1ido_   142    67   3   2   0 SST       Mg    2   65   -1   -1   -1   -1   -1 ...    5 1ido   0 0.1 1.7   ...      |  ZZ                       CEL L ADHESION PR
 69 1c1y_A   17    18   2   2   2 ST         Mg   18   -1   -1   -1   -1   -1   -1 ..     6 1c1y   0 0.2 1.9   ..        | UUZZ   2.7.1.-;        SIG    NALING PROTEI
 70 1ctq_A   17    18   2   2   2 ST         Mg   18   -1   -1   -1   -1   -1   -1 ..     6 1ctq   0 0.1 1.3   ..        | UUZZ                       SIG NALING PROTEI
 71 1f5n_A   52    23   2   2   2 ST         Mg   23   -1   -1   -1   -1   -1   -1 ..     6 1f5n   0 0.2 1.7   ..        | UUZZ                       SIG NALING PROTEI
 72 1cip_A   47   134   2   2   2 ST         Mg 134   -1   -1   -1   -1   -1   -1 ..     6 1cip   0 0.1 1.5   ..        | UUZZ                       HYD ROLASE
 73 1fqj_A   43   134   2   2   2 ST         Mg 134   -1   -1   -1   -1   -1   -1 ..     6 1fqj   1 0.2 2.0   ..        | JUZZ   3.1.4.17;        SIG    NALING PROTEI
 74 1azs_C   54   150   2   2   2 ST         Mg 150   -1   -1   -1   -1   -1   -1 ..     6 1azs   0 0.3 2.3   ..        | UUZZ   4.6.1.1;         COM       PLEX (LYASE/H
 75 1g29_2   43   121   2   0   0 TD         Mg 121   -1   -1   -1   -1   -1   -1 ..     2 1g29   0 0.5 1.9   .b        |                            SUG AR BINDING PR
 76 1ryp_A   17   117   4   1   0 TOOO      Mg 111    3    3   -1   -1   -1   -1 ....  5 1ryp   0 0.3 1.9   ....      |    Z    3.4.99.46;      MUL          TICATALYTIC P
 77 1d0y_A 186    51   2   2   2 TS         Mg   51   -1   -1   -1   -1   -1   -1 ..     6 1d0y   0 0.2 2.0   ..        | JUZZ                       CONTRACTILE PROT
 78 1cxz_A   19    18   2   2   2 TT         Mg   18   -1   -1   -1   -1   -1   -1 ..     6 1cxz   0 0.3 2.2   ..        | UUZZ                       SIG NALING PROTEI
 79 1he1_C   17    18   2   2   2 TT         Mg   18   -1   -1   -1   -1   -1   -1 ..     6 1he1   0 0.2 2.0   ..        | UJZZ                       SIG NALLING COMPL
 80 1tx4_B   19    18   2   2   2 TT         Mg   18   -1   -1   -1   -1   -1   -1 ..     6 1tx4   0 0.2 1.6   ..        | UJZZ                       COM PLEX(GTPASE A
 81 1dek_A   42    66   3   0   0 YQE        Mg   43   23   -1   -1   -1   -1   -1 ...    3 1dek   0 0.6 2.0   ...      |         2.7.4.13;       PHO    SPHOTRANSFERA
```

```
no   cngpname     nsp  np  nw  nn  dons      met  sd1  sd2   sd3  sd4  sd5  sd6  sd7  metal    startaa    |  conf_efimov
  1  8ruc_A  203     1   2   0   3   DE        Mg   -1    1   -1   -1   -1   -1   -1  MG   476  ASPA  203  ..bakb
  2  1bpy_A  190     2   2   1   3  DD         Mg    2   -1   -1   -1   -1   -1   -1  MG   339  ASPA  190  ba.bbbb
  3  1eqo_A   95     2   2   1   3  DD         Mg    2   -1   -1   -1   -1   -1   -1  MG   162  ASPA   95  abbbbba
  4  1uby_   117     4   2   0   1  DD         Mg    4   -1   -1   -1   -1   -1   -1  MG   403  ASP   117  aaaaakkgb
  5  1bl3_B   64    52   2   2   0  DD         Mg   52   -1   -1   -1   -1   -1   -1  MG   503  ASPB   64  bb.bb-bbbkb
  6  1t7p_A  475   179   2   2   1  DD         Mg  179   -1   -1   -1   -1   -1   -1  MG  4002  ASPA  475  bbbbb-d.kbb
  7  1bpy_A  190    66   3   0   1  DDD        Mg    2   64   -1   -1   -1   -1   -1  MG   340  ASPA  190  ba.bbbb-bbbbb
  8  1obw_A   65    37   3   3   0  DDD        Mg    5   32   -1   -1   -1   -1   -1  MG     1  ASPA   65  bakgbbbbbb-jbdbb
  9  1auk_     29   253   4   0   0  DDDN       Mg    1  251    1   -1   -1   -1   -1  MG    44  ASP    29  bbbgbb-bba.bb
 10  1wdc_B   28    11   5   1   0  DDDOD      Mg    2    2    2    5   -1   -1   -1  MG   502  ASPB   28  aadak.kgbbbaaaaa
 11  1vid_   141    29   3   1   2  DDN        Mg   28    1   -1   -1   -1   -1   -1  MG   300  ASP   141  bbd.bakkaaaaaaaaaakgkbbbgbbbbbbgka
 12  1dqn_A  125     1   2   2   2  DE         Mg    1   -1   -1   -1   -1   -1   -1  MG   450  ASPA  125  bbbabb
 13  1gsa_   273     8   2   1   3  DE         Mg    8   -1   -1   -1   -1   -1   -1  MG   319  ASP   273  bbbbb.kbbabbb
 14  1iow_   257    13   2   1   3  DE         Mg   13   -1   -1   -1   -1   -1   -1  MG   331  ASP   257  bbbbbbbkkgbbbba.bb
 15  1qr0_A  107    44   3   1   2  DEE        Mg    2   42   -1   -1   -1   -1   -1  MG   400  ASPA  107  bbbbbbb-aaaaa
 16  1mdl_   195    52   3   1   2  DEE        Mg   26   26   -1   -1   -1   -1   -1  MG   360  ASP   195  bbbkg-bbbga-..kkb
 17  1ec7_A  235    54   3   3   0  DEN        Mg   25   29   -1   -1   -1   -1   -1  MG   498  ASPA  235  bbbkg-.bbga-bbkak
 18  1bx4_A  130     1   2   0   4  DN         Mg    1   -1   -1   -1   -1   -1   -1  MG   365  ASPA  130  b.kbbb
 19  1g97_A  102   125   2   2   2  DN         Mg  125   -1   -1   -1   -1   -1   -1  MG     1  ASPA  102  bkkdk-bbabk
 20  1ig5_A   54     6   4   2   0  DNDO       Mg    2    2    2   -1   -1   -1   -1  MG    78  ASPA   54  aadakgkgbbb
 21  1pox_A  447    29   3   1   2  DNO        Mg   27    2   -1   -1   -1   -1   -1  MG   610  ASPA  447  bbaaaakakaakaaaaaakgbbbbbbbbbkgbba
 22  1pvd_A  444    29   3   1   2  DNO        Mg   27    2   -1   -1   -1   -1   -1  MG   558  ASPA  444  bbaaaaaakaakaaaaaakgbdbbbbbbbkgjda
 23  1zpd_A  440    29   3   1   2  DNO        Mg   27    2   -1   -1   -1   -1   -1  MG   601  ASPA  440  bbaaaakakaakaaaaaakgbdbbbbbbbkgjba
 24  1qs0_A  213    31   3   1   2  DNO        Mg   29    2   -1   -1   -1   -1   -1  MG   501  ASPA  213  .baak-bbkgbb.
 25  1qf5_A   13    27   2   0   5  DO         Mg   27   -1   -1   -1   -1   -1   -1  MG     3  ASPA   13  djgkgaaaaaaakkakbabbbbk.j.kkbbbb
 26  1fiu_B  140    46   2   2   2  DO         Mg   46   -1   -1   -1   -1   -1   -1  MG  3333  ASPB  140  bbabb-bbbba
 27  1t7p_A  475   179   3   0   3  DOD        Mg    1  178   -1   -1   -1   -1   -1  MG  4001  ASPA  475  bbbbbg-d.kbb
 28  1hyo_A  233    24   5   0   0  DOOOT      Mg    1   19    3    1   -1   -1   -1  MG  1004  ASPA  233  bbbbbbaaaaaaakbkkjb.kkakk.bbb
 29  1f1z_A  114    17   3   3   0  DQO        Mg   16    1   -1   -1   -1   -1   -1  MG  2002  ASPA  114  bbabbbbba.babk.bbbbbbb
 30  1ew2_A   42   269   3   3   0  DSE        Mg  113  156   -1   -1   -1   -1   -1  MG  1003  ASPA   42  bbbgd-baaaa-bbbbk
 31  1qb7_A  146     8   2   1   0  DT         Mg    8   -1   -1   -1   -1   -1   -1  MG   335  ASPA  146  bbbabbakgaaaa
```

```
32 5eau_   444     8   3   1   0  DTE        Mg    4    4   -1   -1   -1   -1   -1  MG   733  ASP   444  aaaaaakaaaaak
33 1b0p_A 963    30   3   1   2  DTO        Mg   28    2   -1   -1   -1   -1   -1  MG  1237  ASPA  963  .baaaaaaakaaaaaaaakagbbbbbbbbbkg.ba
34 1g29_2 293    63   3   0   0  DYD        Mg    2   61   -1   -1   -1   -1   -1  MG    16  ASP2  293  akkbbbk-bbbak
35 1g29_2 292     1   2   0   1  ED         Mg    1   -1   -1   -1   -1   -1   -1  MG    79  GLU2  292  bakkbb
36 1a49_B 871    24   2   1   2  ED         Mg   24   -1   -1   -1   -1   -1   -1  MG  1134  GLUB  871  bbk.aaaaakaaaaaaabk.bbbbaaaak
37 1dxe_A 153    26   2   2   1  ED         Mg   26   -1   -1   -1   -1   -1   -1  MG   901  GLUA  153  bbdbaaaaakaaaaakbbgbabbbbbaaaaa
38 1d8c_A 427    28   2   2   2  ED         Mg   28   -1   -1   -1   -1   -1   -1  MG  3001  GLUA  427  ..kbaaaaadaaaaaakdakkbk.bbbbkaaaa
39 1d3y_A 197    52   2   4   0  ED         Mg   52   -1   -1   -1   -1   -1   -1  MG   371  GLUA  197  bbaba-bbbdb
40 1g29_1 292    67   3   0   1  EDK        Mg    1   66   -1   -1   -1   -1   -1  MG    83  GLU1  292  bakkbb-akkbb
41 1qgx_A  72    72   3   1   2  EDO        Mg   70    2   -1   -1   -1   -1   -1  MG   401  GLUA   72  bbabb-bbbabbg
42 1alo_   899     4   2   3   0  EE         Mg    4   -1   -1   -1   -1   -1   -1  MG   918  GLU   899  baaaaaaaa
43 1eyz_A 267    12   2   1   2  EE         Mg   12   -1   -1   -1   -1   -1   -1  MG   402  GLUA  267  bbbbbbb.kbbbbabbb
44 1i7q_A 361   137   2   2   2  EE         Mg  137   -1   -1   -1   -1   -1   -1  MG  1701  GLUA  361  aaaaa-aaaaa
45 1dp0_A 416    45   3   3   0  EHE        Mg    2   43   -1   -1   -1   -1   -1  MG  3001  GLUA  416  dbbdakb-jbgkj
46 1iow_   270     2   2   1   2  EN         Mg    2   -1   -1   -1   -1   -1   -1  MG   330  GLU   270  ba.bbkd
47 1gsa_   281     2   2   1   3  EN         Mg    2   -1   -1   -1   -1   -1   -1  MG   320  GLU   281  babbbka
48 2hgs_A 144   224   3   0   2  ENE        Mg    2  222   -1   -1   -1   -1   -1  MG   501  GLUA  144  ba.bbkb-dbbg.
49 1inp_    79    76   2   0   0  EO         Mg   76   -1   -1   -1   -1   -1   -1  MG   402  GLU    79  b.abb-babbg
50 1ew2_A 216    69   4   1   0  EOED       Mg   53    1   15   -1   -1   -1   -1  MG  1004  GLUA  216  baadk-bbkba.-aabbb
51 1g8f_A  46   120   3   2   0  EOH        Mg  118    2   -1   -1   -1   -1   -1  MG   521  GLUA   46  aaaaa-bbbbbkb
52 1h2r_L  62   490   3   3   0  EOH        Mg  436   54   -1   -1   -1   -1   -1  MG  1005  GLUL   62  gaaaa-bbbbb-aa...
53 1a6o_   161    14   2   0   0  ND         Mg   14   -1   -1   -1   -1   -1   -1  MG   341  ASN   161  aakbbbbaakgbbbbbgka
54 1alo_   649    44   3   1   0  OEO        Mg    2   42   -1   -1   -1   -1   -1  MG   916  ALA   649  bbdbbba-bkjbd
55 1e3m_A  99     3   2   0   4  OO         Mg    3   -1   -1   -1   -1   -1   -1  MG    11  PROA   99  bbakkba.
56 1b25_A  92    89   2   2   2  OO         Mg   89   -1   -1   -1   -1   -1   -1  MG1  800  ASNA   92  bbbbb-.bbbg
57 1bfd_   117     3   3   0   0  OOO        Mg    1    2   -1   -1   -1   -1   -1  MG   531  ASN   117  aakakbkk
58 1ryp_M 183     6   3   1   0  OOO        Mg    3    3   -1   -1   -1   -1   -1  MG     4  THRM  183  aaaakbkkb..
59 1ryp_H 163     6   3   0   0  OOO        Mg    3    3   -1   -1   -1   -1   -1  MG     6  ILEH  163  aaaaabakb.j
60 1ryp_J 166     6   3   2   0  OOO        Mg    3    3   -1   -1   -1   -1   -1  MG     5  ALAJ  166  aaakkbakbjg
61 1ewk_B  89     7   3   2   0  OOO        Mg    6    1   -1   -1   -1   -1   -1  MG  1002  ILEB   89  aaakkbaabgak
62 1ewk_A  89     7   4   2   0  OOOO       Mg    3    3    1   -1   -1   -1   -1  MG  1001  ILEA   89  aaakkbaabgad
63 1dp0_D  15   178   5   0   0  OOOQD      Mg    3    3  142   30   -1   -1   -1  MG  3002  ASPD   15  .dkkdbkkbk.-bbb.k-bbaaa
64 1ezw_A  56     1   2   1   0  OS         Mg    1   -1   -1   -1   -1   -1   -1  MG   361  THRA   56  akbkab
```

```
65 3prn_    28    29   2   3   0 SD          Mg   29   -1   -1   -1   -1   -1   -1 MG   121 SER    28 bbbbbbbbbbbbbbbbbkkgbbbbbbbbbbbbbbgb
66 3pmg_B 116   175   4   1   1 SDDD       Mg 171    2    2   -1   -1   -1   -1 MG     2 SERB 116 bd.kb-bbbakgkbb
67 1qc5_B 342   101   3   3   0 SSD        Mg    2   99   -1   -1   -1   -1   -1 MG   602 SERB 342 bkbaabb-bbk.b
68 1ido_   142    67   3   2   0 SST        Mg    2   65   -1   -1   -1   -1   -1 MG   400 SER   142 bkbakbb-jbbda
69 1c1y_A   17    18   2   2   2 ST         Mg   18   -1   -1   -1   -1   -1   -1 MG   171 SERA   17 gaaaaaaaaaakgbbbabbbbba.
70 1ctq_A   17    18   2   2   2 ST         Mg   18   -1   -1   -1   -1   -1   -1 MG   168 SERA   17 gaaaaaaaaaakgbbbabbdbbab
71 1f5n_A   52    23   2   2   2 ST         Mg   23   -1   -1   -1   -1   -1   -1 MG   595 SERA   52 gaaaaaaaaakgbb..bbbbbkkbb.bbj
72 1cip_A   47   134   2   2   2 ST         Mg 134   -1   -1   -1   -1   -1   -1 MG   356 SERA   47 gaaaa-bbbbj
73 1fqj_A   43   134   2   2   2 ST         Mg 134   -1   -1   -1   -1   -1   -1 MG   352 SERA   43 gaaaa-bbbbj
74 1azs_C   54   150   2   2   2 ST         Mg 150   -1   -1   -1   -1   -1   -1 MG   396 SERC   54 .aaaa-bbbbj
75 1g29_2   43   121   2   0   0 TD         Mg 121   -1   -1   -1   -1   -1   -1 MG   264 THR2   43 gaaaa-bbbgk
76 1ryp_A   17   117   4   1   0 TOOO       Mg 111    3    3   -1   -1   -1   -1 MG     1 THRA   17 abdbb-aaaaabaabbb
77 1d0y_A 186    51   2   2   2 TS         Mg   51   -1   -1   -1   -1   -1   -1 MG   998 THRA 186 gaaaa-dbbkb
78 1cxz_A   19    18   2   2   2 TT         Mg   18   -1   -1   -1   -1   -1   -1 MG   501 THRA   19 gaaaaaaaaaakgbbbabbbbba.
79 1he1_C   17    18   2   2   2 TT         Mg   18   -1   -1   -1   -1   -1   -1 MG   202 THRC   17 gaaaaaaaaaakgbbbabbbbba.
80 1tx4_B   19    18   2   2   2 TT         Mg   18   -1   -1   -1   -1   -1   -1 MG   681 THRB   19 gaaaaaaaaaakg......bbbab
81 1dek_A   42    66   3   0   0 YQE        Mg   43   23   -1   -1   -1   -1   -1 MG   300 TYRA   42 aaaaa-aaaaa-gbbbb
```

```
no   cngpname     nsp  np  nw  nn  dons      met  sd1  sd2  sd3  sd4  sd5  sd6  sd7  his   cn  pdb  cn2  rms  res   carbi    othdon   ecno
  1 1gg1_A   61   265   4   0   1 CHED       Mn  207   34   24   -1   -1   -1   -1 .e..   5 1gg1   0 0.3  2.0   ....     |    U    4.1.2.15;        LYA         SE
  2 1f5a_A 113      2   2   1   2 DD          Mn    2   -1   -1   -1   -1   -1   -1 ..      5 1f5a   3 0.4  2.5   ..       | UUZ    2.7.7.19;        TRA  NSFERASE
  3 1e9g_B 147      5   2   2   2 DD          Mn    5   -1   -1   -1   -1   -1   -1 ..      6 1e9g   0 0.1  1.1   ..       | UUZZ   3.6.1.1;         HYD    ROLASE
  4 1f5a_A 113     54   3   1   2 DDD        Mn    2   52   -1   -1   -1   -1   -1 ...    6 1f5a   0 0.3  2.5   ...      |  BUZ   2.7.7.19;        TRA       NSFERASE
  5 1e9g_B 115     37   3   2   1 DDD        Mn    5   32   -1   -1   -1   -1   -1 ...    6 1e9g   0 0.2  1.1   ...      |  UZZ   3.6.1.1;         HYD       ROLASE
  6 1a6q_    60   222   3   3   0 DDD        Mn  179   43   -1   -1   -1   -1   -1 ...    6 1a6q   1 0.3  2.0   ...      |  ZZZ   3.1.3.16;        HYD        ROLASE
  7 1az9_   260   146   3   2   0 DDE        Mn   11  135   -1   -1   -1   -1   -1 ...    5 1az9   2 0.2  2.0   b..      |  ZZ    3.4.11.9;        PRO       LINE PEPTIDAS
  8 1ga8_A 103   141   3   0   2 DDH        Mn    2  139   -1   -1   -1   -1   -1 ..e    5 1ga8   0 0.1  2.0   .b.      |  UU    2.4.1.44;        TRA          NSFERASE
  9 1i74_B   14   135   4   1   1 DDHD       Mn   61   22   52   -1   -1   -1   -1 ..e.   6 1i74   0 0.2  2.2   ....     |    UZ  3.6.1.1;         HYD         ROLASE
 10 1qmg_A 315      4   2   2   2 DE          Mn    4   -1   -1   -1   -1   -1   -1 ..      6 1qmg   0 0.1  1.6   ..       | UUZZ   1.1.1.86;        OXI       DOREDUCTASE
 11 1f3i_A   97   229   2   2   1 DE         Mn  229   -1   -1   -1   -1   -1   -1 ..      5 1f3i   1 0.4  2.3   ..       | UZZ                        TRANSCRIPTION/DN
 12 1muc_A 198     51   3   3   0 DED        Mn   26   25   -1   -1   -1   -1   -1 ...    6 1muc   0 0.3  1.9   ...      |  ZZZ   5.5.1.1          ISO        MERASE
 13 1hhs_A 454     41   3   3   0 DEO        Mn   37    4   -1   -1   -1   -1   -1 ...    6 1hhs   0 0.2  2.0   ...      |  ZZZ                       RNA  POLYMERASE
 14 1d3v_A 124   110   4   0   2 DHDD      Mn    2  106    2   -1   -1   -1   -1 .d..   6 1d3v   1 0.3  1.7   ...b     |    UU  3.5.3.1;         HYD             ROLASE
 15 1ii7_A    8   200   4   1   1 DHDH       Mn    2   39  159   -1   -1   -1   -1 .e.e   6 1ii7   0 0.2  2.2   ....     |    UZ                      REP LICATION
 16 1az9_   271   135   4   2   0 DHEE       Mn   83   29   23   -1   -1   -1   -1 .e..   6 1az9   1 0.3  2.0   ....     |    ZZ  3.4.11.9;        PRO          LINE PEPTIDAS
 17 1eqj_A 403     59   3   0   2 DHH        Mn    4   55   -1   -1   -1   -1   -1 .ee    5 1eqj   0 0.3  1.7   b..      |  UU    5.4.2.1;         ISO      MERASE
 18 1ii7_A   49   157   4   1   1 DNHH       Mn   35   89   33   -1   -1   -1   -1 ..ed   6 1ii7   0 0.2  2.2   ....     |    UZ                      REP  LICATION
 19 1a6q_    60      1   2   4   0 DO          Mn    1   -1   -1   -1   -1   -1   -1 ..      6 1a6q   0 0.3  2.0   ..       | ZZZZ   3.1.3.16;        HYD   ROLASE
 20 3bam_A   94     18   2   1   1 DO         Mn   18   -1   -1   -1   -1   -1   -1 ..      4 3bam   1 0.3  1.8   ..       | UZ     3.1.21.4;        HYD      ROLASE/DNA
 21 1ksi_A 451   142   5   1   0 DODDO     Mn    1    1  139    1   -1   -1   -1 ..... 6 1ksi   0 0.3  2.2   .....    |     Z  1.4.3.6;         OXI        DOREDUCTASE
 22 1eqj_A   12   433   4   0   0 DSDH       Mn   50  382    1   -1   -1   -1   -1 ...e   4 1eqj   1 0.3  1.7   b...     |        5.4.2.1;         ISO     MERASE
 23 1nls_     8     16   4   2   0 EDDH       Mn    2    9    5   -1   -1   -1   -1 ...e   6 1nls   0 0.1  0.9   ....     |    ZZ                      LUAGGTININ
 24 1f3i_A 110   235   2   3   0 EE         Mn  235   -1   -1   -1   -1   -1   -1 ..      5 1f3i   1 0.3  2.3   b.       | ZZZ                        TRANSCRIPTION/DN
 25 1d8h_B 305   191   3   2   0 EEE        Mn    2  189   -1   -1   -1   -1   -1 ...    5 1d8h   1 0.4  2.0   ...      |  ZZ    3.1.3.33;        HYD         ROLASE
 26 1f52_A 131     89   3   2   0 EEE        Mn   81    8   -1   -1   -1   -1   -1 ...    5 1f52   1 0.2  2.5   ...      |  ZZ    6.3.1.2;         LIG   ASE
 27 1f52_A 129   228   3   2   2 EHE        Mn  140   88   -1   -1   -1   -1   -1 .d.    7 1f52   0 0.3  2.5   ...      |  UUZZ 6.3.1.2;         LIG            ASE
 28 1a9x_A 299      2   2   1   2 EN          Mn    2   -1   -1   -1   -1   -1   -1 ..      5 1a9x   0 0.2  1.8   b.       | UUZ                        ADDITRANSFERASE
 29 1cpo_   104      4   3   2   1 EOS         Mn    1    3   -1   -1   -1   -1   -1 . .    6 1cpo   0 0.2  1.9   ...      |  UZZ   1.11.1.10        OXI   DOREDUCTASE
 30 1b8a_A 361      3   2   2   3 ES          Mn    3   -1   -1   -1   -1   -1   -1 ..      7 1b8a   0 0.4  1.9   ..       | JUUZZ 6.1.1.12;        LIG     ASE
 31 1i74_B    8     67   3   1   1 HDD         Mn    4   63   -1   -1   -1   -1   -1 e..    5 1i74   1 0.2  2.2   .b.      |  UZ    3.6.1.1;         HYD    ROLASE
```

```
32 1d3v_A 101   131   4   0   3 HDDD       Mn   23    4 104   -1   -1   -1   -1 d...   7 1d3v   0 0.3 1.7   ....     |    JUU 3.5.3.1;         HYD              ROLASE
33 1i0h_A   26   145   4   0   1 HHDH       Mn   55   86    4   -1   -1   -1   -1 ee.e   5 1i0h   0 0.1 1.4   ....     |    U    1.15.1.1;        OXI         DOREDUCTASE
34 1fi2_A   88    49   4   2   0 HHEH       Mn    2    5   42   -1   -1   -1   -1 ee.e   6 1fi2   0 0.2 1.6   ....     |    ZZ   1.2.3.4           OXI     DOREDUCTASE
35 1a9x_A 829    12   2   0   2 QE          Mn   12   -1   -1   -1   -1   -1   -1 ..     4 1a9x   0 0.2 1.8   ..        | UU                              AMIDOTRANSFERASE
36 1a9x_A 285    14   2   0   3 QE          Mn   14   -1   -1   -1   -1   -1   -1 ..     5 1a9x   0 0.2 1.8   ..        | UUU                             AMIDOTRANSFERASE
37 1a9x_C2829    12   3   1   2 QQE        Mn    0   12   -1   -1   -1   -1   -1 ...    6 1a9x   0 0.4 1.8   ...      |   UUZ                           AMI  DOTRANSFERASE
```

```
no   cngpname     nsp  np  nw  nn  dons     met  sd1  sd2   sd3  sd4  sd5  sd6  sd7  metal    startaa    |  conf_efimov
   1 1gg1_A   61   265   4   0   1  CHED      Mn  207   34   24   -1   -1   -1   -1 MN   371  CYSA   61  .babb-kb.ka-bbbbb-kabbb
   2 1f5a_A  113     2   2   1   2  DD        Mn    2   -1   -1   -1   -1   -1   -1 MN  1002  ASPA  113  gbbbbbb
   3 1e9g_B  147     5   2   2   2  DD        Mn    5   -1   -1   -1   -1   -1   -1 MN  2008  ASPB  147  bbbggbbdbb
   4 1f5a_A  113    54   3   1   2  DDD       Mn    2   52   -1   -1   -1   -1   -1 MN  1001  ASPA  113  gbbbbbb-bbbbb
   5 1e9g_B  115    37   3   2   1  DDD       Mn    5   32   -1   -1   -1   -1   -1 MN  2005  ASPB  115  bbkgbbbbbb-bbdbb
   6 1a6q_    60   222   3   3   0  DDD       Mn  179   43   -1   -1   -1   -1   -1 MN     1  ASP    60  bbbbb-bbkaa-bdbbb
   7 1az9_   260   146   3   2   0  DDE       Mn   11  135   -1   -1   -1   -1   -1 MN   442  ASP   260  bbbbb-bbbbb-bbbbb
   8 1ga8_A  103   141   3   0   2  DDH       Mn    2  139   -1   -1   -1   -1   -1 MN   400  ASPA  103  bbbadbb-bbbdb
   9 1i74_B   14   135   4   1   1  DDHD      Mn   61   22   52   -1   -1   -1   -1 MN   402  ASPB   14  baaaa-bbg..-bbkbb-aaakg
  10 1qmg_A  315     4   2   2   2  DE        Mn    4   -1   -1   -1   -1   -1   -1 MN   601  ASPA  315  aaaaaaaa.
  11 1f3i_A   97   229   2   2   1  DE        Mn  229   -1   -1   -1   -1   -1   -1 MN   498  ASPA   97  bb.bb-akaaa
  12 1muc_A  198    51   3   3   0  DED       Mn   26   25   -1   -1   -1   -1   -1 MN   373  ASPA  198  bbbkg-.bbga-bb.ak
  13 1hhs_A  454    41   3   3   0  DEO       Mn   37    4   -1   -1   -1   -1   -1 MN   665  ASPA  454  .kb.b-bbbbgbbbg
  14 1d3v_A  124   110   4   0   2  DHDD      Mn    2  106    2   -1   -1   -1   -1 MN   501  ASPA  124  bb.kbkg-bbbakab
  15 1ii7_A    8   200   4   1   1  DHDH      Mn    2   39  159   -1   -1   -1   -1 MN   403  ASPA    8  b.gagbg-djgab-.babb
  16 1az9_   271   135   4   2   0  DHEE      Mn   83   29   23   -1   -1   -1   -1 MN   441  ASP   271  bbbbb-abbbb-bbdb.-bbbbb
  17 1eqj_A  403    59   3   0   2  DHH       Mn    4   55   -1   -1   -1   -1   -1 MN   601  ASPA  403  aaaakakkg-kbbbk
  18 1ii7_A   49   157   4   1   1  DNHH      Mn   35   89   33   -1   -1   -1   -1 MN   404  ASPA   49  djgab-bbaak-bbbbb-bb.ba
  19 1a6q_    60     1   2   4   0  DO        Mn    1   -1   -1   -1   -1   -1   -1 MN     2  ASP    60  bbbbbg
  20 3bam_A   94    18   2   1   1  DO        Mn   18   -1   -1   -1   -1   -1   -1 MN   901  ASPA   94  bbabb.bbbd.gbbbbbbbbbbk
  21 1ksi_A  451   142   5   1   0  DODDO     Mn    1    1  139    1   -1   -1   -1 MN   653  ASPA  451  bbbbdbg-gbbbbb
  22 1eqj_A   12   433   4   0   0  DSDH      Mn   50  382    1   -1   -1   -1   -1 MN   701  ASPA   12  bbbgb-b.aaa-.bk.b.
  23 1nls_     8    16   4   2   0  EDDH      Mn    2    9    5   -1   -1   -1   -1 MN   239  GLU     8  bbbbbkbbbakkgbbd.bbbb
  24 1f3i_A  110   235   2   3   0  EE        Mn  235   -1   -1   -1   -1   -1   -1 MN   499  GLUA  110  aakbb-bkbaa
  25 1d8h_B  305   191   3   2   0  EEE       Mn    2  189   -1   -1   -1   -1   -1 MN   502  GLUB  305  kbbbbbb-bbbbb
  26 1f52_A  131    89   3   2   0  EEE       Mn   81    8   -1   -1   -1   -1   -1 MN   469  GLUA  131  dbbbb-.bbab-bbbbb
  27 1f52_A  129   228   3   2   2  EHE       Mn  140   88   -1   -1   -1   -1   -1 MN   470  GLUA  129  bbdbb-bdbbb-.bbbk
  28 1a9x_A  299     2   2   1   2  EN        Mn    2   -1   -1   -1   -1   -1   -1 MN  1901  GLUA  299  ba.bbd.
  29 1cpo_   104     4   3   2   1  EOS       Mn    1    3   -1   -1   -1   -1   -1 MN   301  GLU   104  .abbbgbak
  30 1b8a_A  361     3   2   2   3  ES        Mn    3   -1   -1   -1   -1   -1   -1 MN   601  GLUA  361  gbbkbb..
  31 1i74_B    8    67   3   1   1  HDD       Mn    4   63   -1   -1   -1   -1   -1 MN   401  HISB    8  bbbbgbbaa-bbg..
```

```
32 1d3v_A 101   131   4   0   3 HDDD       Mn   23    4 104   -1   -1   -1   -1 MN   500 HISA 101 .bkkk-bb.kbkgkb-bbbak
33 1i0h_A   26   145   4   0   1 HHDH       Mn   55   86    4   -1   -1   -1   -1 MN   206 HISA   26 aaaaa-aaaaa-.bbkbaakk
34 1fi2_A   88    49   4   2   0 HHEH       Mn    2    5   42   -1   -1   -1   -1 MN   202 HISA   88 bbbbbak.bbbb-bbbbb
35 1a9x_A 829    12   2   0   2 QE          Mn   12   -1   -1   -1   -1   -1   -1 MN 1911 GLNA 829 bbbbbbbggbbbba.bb
36 1a9x_A 285    14   2   0   3 QE          Mn   14   -1   -1   -1   -1   -1   -1 MN 1902 GLNA 285 bbbbbbbka.gbbbba.bb
37 1a9x_C2829    12   3   1   2 QQE        Mn    0   12   -1   -1   -1   -1   -1 MN 3911 GLNC2829 bbbbbbbggbbbba.bb
```

```
no   cngpname     nsp  np  nw  nn  dons      met  sd1  sd2  sd3  sd4  sd5  sd6  sd7  his   cn  pdb  cn2  rms  res   carbi    othdon   ecno
   1  1eg9_A   81    20   2   0   3  CC         Fe   20   -1   -1   -1   -1   -1   -1  ..      5  1eg9   0  0.3  1.6   ..        |  JXX    1.14.12.12       OXI    DOREDUCTASE
   2  1h2r_L   84   465   2   0   2  CC         Fe  465   -1   -1   -1   -1   -1   -1  ..      4  1h2r   2  0.2  1.4   ..        |  JX     1.12.2.1;        OXI  DOREDUCTASE
   3  1b71_A  158    19   4   0   0  CCCC       Fe    3   13    3   -1   -1   -1   -1  ....   4  1b71   0  0.1  1.9   ....     |                                ELECTRON  TRANSPO
   4  1dxg_A    9    20   4   0   0  CCCC       Fe    3   16    1   -1   -1   -1   -1  ....   4  1dxg  20  0.1  1.8   ....     |                                NON-HEME  IRON  PR
   5  1rb9_     6    36   4   0   0  CCCC       Fe    3   30    3   -1   -1   -1   -1  ....   4  1rb9   0  0.1  0.9   ....     |                             IRON-SULFUR PROT
   6  1e5d_A   83   321   4   1   1  DDHO       Fe   82   61  178   -1   -1   -1   -1  ..e.   6  1e5d   1  0.3  2.5   ....     |    UZ                          OXI   DOREDUCTASE
   7  1ute_A   14   487   5   0   1  DDYH  O    Fe   38    3  168   -1  278   -1   -1  ...e  .6  1ute   0  0.2  1.5   ....  .  |    U    3.1.3.2;         HYD           ROLASE
   8  1r2f_B   67   125   4   0   0  DEHE       Fe   31    3   91   -1   -1   -1   -1  ..d.   4  1r2f   0  0.4  2.1   ...b     |          1.17.4.1;        OXI   DOREDUCTASE
   9  1ute_A   52   449   5   0   1  DNHH  O    Fe   39   95   35   -1  280   -1   -1  ..ed  .6  1ute   0  0.3  1.5   ....  .  |    U    3.1.3.2;         HYD            ROLASE
  10  1b71_A   53    78   4   1   0  EEEH       Fe   41   34    3   -1   -1   -1   -1  ...d   5  1b71   0  0.1  1.9   .b..     |    Z                          ELECTRON  TRANSPO
  11  1afr_A  143    89   4   0   0  EEEH       Fe   53   33    3   -1   -1   -1   -1  ...d   4  1afr   0  0.3  2.4   .b..     |          1.14.99.6;       OXI    DOREDUCTASE
  12  1r2f_B   98    97   4   0   0  EEEH       Fe   60   34    3   -1   -1   -1   -1  ...d   4  1r2f   0  0.3  2.1   .b..     |          1.17.4.1;        OXI  DOREDUCTASE
  13  1mty_E  144   102   4   2   0  EEEH       Fe   65   34    3   -1   -1   -1   -1  ...d   6  1mty   1  0.3  1.7   ....     |    ZZ  1.14.13.25       MON              OOXYGENASE
  14  1mty_E  114    33   3   3   0  EEH        Fe   30    3   -1   -1   -1   -1   -1  ..d    6  1mty   1  0.3  1.7   ...      |  ZZZ   1.14.13.25       MON           OOXYGENASE
  15  1afr_A  105   124   4   0   0  EEHE       Fe   38    3   83   -1   -1   -1   -1  ..d.   4  1afr   0  0.3  2.4   b...     |          1.14.99.6;       OXI      DOREDUCTASE
  16  1guq_B  182   116   4   0   0  EHHH       Fe   99   15    2   -1   -1   -1   -1  .dee   4  1guq   0  0.1  1.8   b...     |          2.7.7.10;        NUC           LEOTIDYLTRANS
  17  1dqi_A   14   100   6   0   0  EHHHCH     Fe    2   25    6   64    3   -1   -1  .eee.d6  1dqi   0  0.2  1.7   ......  |                                OXI   DOREDUCTASE
  18  2hmq_A   58    48   5   0   1  EHHHD      Fe   15    4   24    5   -1   -1   -1  .eee.  6  2hmq   0  0.2  1.7   .....    |     U                           OXY      GEN  TRANSPORT
  19  1bk0_   214    56   3   1   1  HDH        Fe    2   54   -1   -1   -1   -1   -1  e.e     5  1bk0   0  0.2  1.3   ...      |  XZ                              LACTAM  ANTIBIO
  20  1ds1_A  144   135   3   1   2  HEH        Fe    2  133   -1   -1   -1   -1   -1  e.e    6  1ds1   0  0.1  1.1   ...      |  UUZ                          OXI   DOREDUCTASE,L
  21  1e5d_B   79   325   5   0   0  HEHDO      Fe    2   65   19  239   -1   -1   -1  e.e..  5  1e5d   0  0.2  2.5   .....   |                                OXI   DOREDUCTASE
  22  1mrp_     9   187   4   1   1  HEYY       Fe   48  138    1   -1   -1   -1   -1  e...    6  1mrp   0  0.1  1.6   ....     |    UZ                          IRON  TRANSPORT
  23  1eg9_A   83    21   2   0   3  HH         Fe   21   -1   -1   -1   -1   -1   -1  dd      5  1eg9   0  0.3  1.6   ..        |  JXX    1.14.12.12       OXI       DOREDUCTASE
  24  1dxr_C  124   185   2   0   4  HH         Fe  185   -1   -1   -1   -1   -1   -1  ee      6  1dxr   0  0.1  2.0   ..       |  BBBB                          PHO   TOSYNTHETIC  R
  25  1eg9_A  208   154   3   0   1  HHD        Fe    5  149   -1   -1   -1   -1   -1  ee.     4  1eg9   0  0.2  1.6   ..b      |   U     1.14.12.12       OXI          DOREDUCTASE
  26  4pah_   285    45   3   1   2  HHE        Fe    5   40   -1   -1   -1   -1   -1  ee.     6  4pah   0  0.2  2.0   ...      |   UUZ   1.14.16.1;       MON         OOXGYGENASE
  27  1bou_D   12   230   3   1   0  HHE        Fe   49  181   -1   -1   -1   -1   -1  ee.     4  1bou   1  0.1  2.2   ...      |   Z     1.13.11.8;       DIO        XYGENASE
  28  1cjx_D  161   161   3   0   1  HHE        Fe   79   82   -1   -1   -1   -1   -1  ee.     4  1cjx   1  0.1  2.4   ...      |   U     1.13.11.27       OXI        DOREDUCTASE
  29  2hmq_D   25    81   4   0   1  HHED       Fe   29    4   48   -1   -1   -1   -1  ee..   5  2hmq   1  0.2  1.7   .....    |     U                           OXY   GEN  TRANSPORT
  30  1cjx_A  161   469   4   0   1  HHEO       Fe   79   82  308   -1   -1   -1   -1  ee..   5  1cjx   0  0.3  2.4   ....     |    U    1.13.11.27       OXI           DOREDUCTASE
  31  1yge_   499   191   3   1   1  HHH        Fe    5  186   -1   -1   -1   -1   -1  eee     5  1yge   1  0.3  1.4   ...      |   UZ    1.13.11.12       DIO          XYGENASE
```

```
32 1dqi_B   16    98   5   1   0  HHHCH     Fe   25    6   64    3   -1   -1   -1  eee.d  6  1dqi   0  0.3  1.7   .....    |     Z                        OXI DOREDUCTASE
33 1lox_   361   184   4   0   1  HHHH      Fe    5  175    4   -1   -1   -1   -1  eeed   5  1lox   1  0.2  2.4   ....     |    U    1.13.11.33       OXI          DOREDUCTASE
34 1dxr_C  233    15   2   0   4  MH         Fe   15   -1   -1   -1   -1   -1   -1  .e      6  1dxr   0  0.1  2.0   ..        |  BBBB                      PHO TOSYNTHETIC  R
35 1dxr_C   74    17   2   0   4  MH         Fe   17   -1   -1   -1   -1   -1   -1  .e      6  1dxr   0  0.1  2.0   ..        |  BBBB                      PHOTOSYNTHETIC  R
36 1dxr_C  110    26   2   0   4  MH         Fe   26   -1   -1   -1   -1   -1   -1  .e      6  1dxr   0  0.1  2.0   ..        |  BBBB                      PHO TOSYNTHETIC  R
37 3pcg_O  408    54   4   0   1  YYHH      Fe   39   13    2   -1   -1   -1   -1  ..ee   5  3pcg   0  0.2  2.0   ....     |    U    1.13.11.3;       DIO          XYGENASE
38 2ahj_C  109     4   3   0   4  CC#        Fe    3    1   -1   -1   -1   -1   -1  . .     6  2ahj   0  0.2  1.7   . .       |X  BXB   4.2.1.84;        LYA    SE
```

```
no   cngpname     nsp  np  nw  nn  dons      met  sd1  sd2   sd3  sd4  sd5  sd6  sd7  metal    startaa    |  conf_efimov

   1  1eg9_A   81    20   2   0   3  CC         Fe   20   -1   -1   -1   -1   -1   -1  FE2  751  CYSA   81  dbbkakgbbabkba..bbk.bbbka

   2  1h2r_L   84   465   2   0   2  CC         Fe  465   -1   -1   -1   -1   -1   -1  FE  1004  CYSL   84  kadaa-aaaaa

   3  1b71_A  158    19   4   0   0  CCCC       Fe    3   13    3   -1   -1   -1   -1  FE   192  CYSA  158  bbbaakgbbbb.abbbk.baakgb

   4  1dxg_A    9    20   4   0   0  CCCC       Fe    3   16    1   -1   -1   -1   -1  FE    37  CYSA    9  bbbaakgbbbbbabbbkbbbbbggb

   5  1rb9_     6    36   4   0   0  CCCC       Fe    3   30    3   -1   -1   -1   -1  FE    54  CYS     6  bbbkakgb-bbbkakgb

   6  1e5d_A   83   321   4   1   1  DDHO       Fe   82   61  178   -1   -1   -1   -1  FE2  404  ASPA   83  bkkak-k.abg-babjb-.....

   7  1ute_A   14   487   5   0   1  DDYH  O    Fe   38    3  168   -1  278   -1   -1  FE1  501  ASPA   14  bbgkk-kjgabbkd-.bd.b

   8  1r2f_B   67   125   4   0   0  DEHE       Fe   31    3   91   -1   -1   -1   -1  FE   402  ASPB   67  aaaaa-aaaaaaaa-aaaaa

   9  1ute_A   52   449   5   0   1  DNHH  O    Fe   39   95   35   -1  280   -1   -1  FE2  501  ASPA   52  kjgab-bbaaa-bbabb-b..bd

  10  1b71_A   53    78   4   1   0  EEEH       Fe   41   34    3   -1   -1   -1   -1  FE   193  GLUA   53  aaaaa-aaaaa-aaaaaaaa

  11  1afr_A  143    89   4   0   0  EEEH       Fe   53   33    3   -1   -1   -1   -1  FE   364  GLUA  143  aaaaa-aaaaa-aaaaaaaa

  12  1r2f_B   98    97   4   0   0  EEEH       Fe   60   34    3   -1   -1   -1   -1  FE   403  GLUB   98  aaaaa-aaak.-aaaaaaaa

  13  1mty_E  144   102   4   2   0  EEEH       Fe   65   34    3   -1   -1   -1   -1  FE     1  GLUE  144  aaaaa-aaaaa-kaaaaaaa

  14  1mty_E  114    33   3   3   0  EEH        Fe   30    3   -1   -1   -1   -1   -1  FE     2  GLUE  114  aaaaa-aaaaaaaa

  15  1afr_A  105   124   4   0   0  EEHE       Fe   38    3   83   -1   -1   -1   -1  FE   365  GLUA  105  aaakk-aaaaaaaa-aaaaa

  16  1guq_B  182   116   4   0   0  EHHH       Fe   99   15    2   -1   -1   -1   -1  FE   351  GLUB  182  aaaaa-.bbbb-bb.bbbb

  17  1dqi_A   14   100   6   0   0  EHHHCH     Fe    2   25    6   64    3   -1   -1  FE   501  GLUA   14  b.bbabb-bbbbdbaab.b-bbbakkjb

  18  2hmq_A   58    48   5   0   1  EHHHD      Fe   15    4   24    5   -1   -1   -1  FE1  101  GLUA   58  aaaaa-aaaaaaaaa-aaaaaaaaak

  19  1bk0_   214    56   3   1   1  HDH        Fe    2   54   -1   -1   -1   -1   -1  FE   350  HIS   214  bb.bbb.-bbbbb

  20  1ds1_A  144   135   3   1   2  HEH        Fe    2  133   -1   -1   -1   -1   -1  FE   341  HISA  144  bbbbbga-bbbbb

  21  1e5d_B   79   325   5   0   0  HEHDO      Fe    2   65   19  239   -1   -1   -1  FE1  404  HISB   79  bk.kbkk-gdkba-k.abg-.....

  22  1mrp_     9   187   4   1   1  HEYY       Fe   48  138    1   -1   -1   -1   -1  FE   310  HIS     9  kbbaa-bbaba-baaaaa

  23  1eg9_A   83    21   2   0   3  HH         Fe   21   -1   -1   -1   -1   -1   -1  FE1  751  HISA   83  bkakgbbabkba..bbk.bbbkakgb

  24  1dxr_C  124   185   2   0   4  HH         Fe  185   -1   -1   -1   -1   -1   -1  FE   404  HISC  124  akak.-aaagg

  25  1eg9_A  208   154   3   0   1  HHD        Fe    5  149   -1   -1   -1   -1   -1  FE   752  HISA  208  aakaaaakaa-aakaa

  26  4pah_   285    45   3   1   2  HHE        Fe    5   40   -1   -1   -1   -1   -1  FE   425  HIS   285  aaaaaaaaaa-aaakb

  27  1bou_D   12   230   3   1   0  HHE        Fe   49  181   -1   -1   -1   -1   -1  FE   502  HISD   12  .bkba-bbdga-kkaaa

  28  1cjx_D  161   161   3   0   1  HHE        Fe   79   82   -1   -1   -1   -1   -1  FE   629  HISD  161  babbb-babbb-bbbbb

  29  2hmq_D   25    81   4   0   1  HHED       Fe   29    4   48   -1   -1   -1   -1  FE2  101  HISD   25  aaaaa-aaaaaaaaa-aaaaa

  30  1cjx_A  161   469   4   0   1  HHEO       Fe   79   82  308   -1   -1   -1   -1  FE   629  HISA  161  ba.bb-babbb-bbbbb-.....

  31  1yge_   499   191   3   1   1  HHH        Fe    5  186   -1   -1   -1   -1   -1  FE   840  HIS   499  aaaaaaaaaa-aaaaa
```

```
32 1dqi_B   16    98   5   1   0 HHHCH     Fe   25    6   64    3   -1   -1   -1 FE   502 HISB   16 ..abb-bbbbdbaab.b-bbbakkjb
33 1lox_   361   184   4   0   1 HHHH      Fe    5 175    4   -1   -1   -1   -1 FE   840 HIS   361 aaaaaaaaaa-aaaaaaaak
34 1dxr_C 233    15   2   0   4 MH         Fe   15   -1   -1   -1   -1   -1   -1 FE   403 METC 233 aaaaaaaaakgbbakaabbk
35 1dxr_C   74    17   2   0   4 MH         Fe   17   -1   -1   -1   -1   -1   -1 FE   401 METC   74 aaaaaaaaaadaaajakaabba
36 1dxr_C 110    26   2   0   4 MH         Fe   26   -1   -1   -1   -1   -1   -1 FE   402 METC 110 aaaaaaaaaaaakdakak.kajbbaaaaagg
37 3pcg_O 408    54   4   0   1 YYHH      Fe   39   13    2   -1   -1   -1   -1 FE   600 TYRO 408 gbbdb-bbbbb-bbbbbbb
38 2ahj_C 109     4   3   0   4 CC#        Fe    3    1   -1   -1   -1   -1   -1 FE   300 CYSC 109 bbba.....
```

```
no   cngpname     nsp  np  nw  nn  dons     met  sd1  sd2  sd3  sd4  sd5  sd6  sd7  his   cn  pdb  cn2  rms  res   carbi    othdon   ecno
   1 1oac_A  524   165   3   0   0   HHH      Cu   -1    2  163   -1   -1   -1   -1  .eed   4  1oac   0  0.1  2.0   ....    |         1.4.3.6          OXI    DOREDUCTASE
   2 1ehk_B  149     8   3   0   1  CCH       Cu    4    4   -1   -1   -1   -1   -1  ..d    4  1ehk   1  0.3  2.4   ...     |   J     1.9.3.1;         OXIDOREDUCTASE
   3 1cyx_   207     8   4   0   1  COCH      Cu    2    2    4   -1   -1   -1   -1  ...d   5  1cyx   0  0.3  2.3   ....    |   J    1.10.3.-;        ELECTRON  TRANSPO
   4 2cua_A  149     8   4   0   1  COCH      Cu    2    2    4   -1   -1   -1   -1  ...d   5  2cua   0  0.4  1.6   ....    |   J    1.9.3.1;         ELE CTRON  TRANSPO
   5 2occ_B  196     8   4   0   1  COCH      Cu    2    2    4   -1   -1   -1   -1  ...d   5  2occ   0  0.2  2.3   ....    |   J    1.9.3.1;         OXI DOREDUCTASE
   6 1cyx_   172    39   3   0   1  HCC       Cu   35    4   -1   -1   -1   -1   -1  d..    4  1cyx   1  0.2  2.3   ...     |   J     1.10.3.-;        ELECTRON  TRANSPO
   7 1ehk_B  114    46   4   0   1  HCCM      Cu   35    4    7   -1   -1   -1   -1  d...   5  1ehk   0  0.3  2.4   ....    |   J    1.9.3.1;         OXI    DOREDUCTASE
   8 2cua_A  114    46   4   0   1  HCCM      Cu   35    4    7   -1   -1   -1   -1  d...   5  2cua   0  0.3  1.6   ....    |   J    1.9.3.1;         ELE    CTRON  TRANSPO
   9 2occ_B  161    46   4   0   1  HCCM      Cu   35    4    7   -1   -1   -1   -1  d...   5  2occ   0  0.3  2.3   ....    |   J    1.9.3.1;         OXI    DOREDUCTASE
  10 2mta_A   53    42   3   0   0  HCH       Cu   39    3   -1   -1   -1   -1   -1  d.d    3  2mta   1  0.3  2.4   ...     |         1.4.99.3)        ELE    CTRON  TRANSPO
  11 1a65_A  396    61   3   0   0  HCH       Cu   56    5   -1   -1   -1   -1   -1  d.d    3  1a65   0  0.2  2.2   ...     |         1.10.3.2;        OXI    DOREDUCTASE
  12 1pmy_    40    46   4   0   0  HCHM      Cu   38    3    5   -1   -1   -1   -1  d.d.   4  1pmy   0  0.3  1.5   ....    |                             ELECTRON  TRANSFE
  13 1f56_A   34    50   4   0   0  HCHM      Cu   40    5    5   -1   -1   -1   -1  d.d.   4  1f56   0  0.3  2.0   ....    |                             PLANT PROTEIN
  14 1plc_    37    55   4   0   0  HCHM      Cu   47    3    5   -1   -1   -1   -1  d.d.   4  1plc   0  0.4  1.3   ....    |                           ELECTRON TRANSPO
  15 1aoz_A  445    72   4   0   0  HCHM      Cu   62    5    5   -1   -1   -1   -1  d.d.   4  1aoz   0  0.4  1.9   ....    |         1.10.3.3)        OXI      DOREDUCTASE(O
  16 1qhq_A   57    75   4   0   0  HCHM      Cu   65    5    5   -1   -1   -1   -1  d.d.   4  1qhq   0  0.4  1.5   ....    |                             ELECTRON  TRANSFE
  17 1jer_    46    53   4   0   0  HCHQ      Cu   43    5    5   -1   -1   -1   -1  d.d.   4  1jer  20  0.2  1.6   ....    |                          ELECTRONE TRANSPO
  18 1e30_A   85    63   4   0   0  HCHQ      Cu   53    5    5   -1   -1   -1   -1  d.d.   4  1e30   0  0.2  1.5   ....    |                             RUSTICYANIN
  19 1bxa_A   53    45   3   0   0  HCM       Cu   39    6   -1   -1   -1   -1   -1  d..    3  1bxa   0  0.5  1.3   ...     |                            ELECTRON  TRANSPO
  20 1aoz_A   60   388   2   0   1  HH        Cu  388   -1   -1   -1   -1   -1   -1  ee     3  1aoz   0  0.1  1.9   ..      | U       1.10.3.3)        OXI     DOREDUCTASE(O
  21 1phm_   107    65   3   0   0  HHH       Cu    1   64   -1   -1   -1   -1   -1  ddd    3  1phm   0  0.2  1.9   ...     |         1.14.17.3;       MON         OOXYGENASE
  22 1yai_B   45    80   3   0   0  HHH       Cu    2   78   -1   -1   -1   -1   -1  dee    3  1yai   1  0.1  1.9   ...     |         1.15.1.1;        OXI  DOREDUCTASE
  23 1ksi_A  442   161   3   2   0  HHH       Cu    2  159   -1   -1   -1   -1   -1  eed    5  1ksi   0  0.3  2.2   ...     |  ZZ    1.4.3.6;         OXI       DOREDUCTASE
  24 1lla_   173    31   3   0   0  HHH       Cu    4   27   -1   -1   -1   -1   -1  eee    3  1lla  20  0.1  2.2   ...     |                            OXYGEN TRANSPORT
  25 1bt3_A  240    34   3   0   2  HHH       Cu    4   30   -1   -1   -1   -1   -1  eee    5  1bt3   0  0.4  2.5   ...     |  JU    1.10.3.1;        OXI      DOREDUCTASE
  26 1lla_   324    40   3   0   0  HHH       Cu    4   36   -1   -1   -1   -1   -1  eee    3  1lla  20  0.2  2.2   ...     |                            OXYGEN TRANSPORT
  27 1bt3_A   88    30   3   0   2  HHH       Cu   21    9   -1   -1   -1   -1   -1  eee    5  1bt3   0  0.4  2.5   ...     |  JU    1.10.3.1;        OXI     DOREDUCTASE
  28 1aoz_A   62   446   3   0   1  HHH       Cu   42  404   -1   -1   -1   -1   -1  dee    4  1aoz   0  0.2  1.9   ...     |  U     1.10.3.3)        OXI        DOREDUCTASE(O
  29 1a65_A   66   387   3   0   0  HHH       Cu   43  344   -1   -1   -1   -1   -1  dee    3  1a65   0  0.2  2.2   ...     |        1.10.3.2;        OXI       DOREDUCTASE
  30 1ehk_A  233    50   3   0   1  HHH       Cu   49    1   -1   -1   -1   -1   -1  dee    4  1ehk   0  0.3  2.4   ...     |  U     1.9.3.1;         OXI     DOREDUCTASE
  31 2occ_A  240    51   3   0   1  HHH       Cu   50    1   -1   -1   -1   -1   -1  dee    4  2occ   1  0.2  2.3   ...     |  U     1.9.3.1;         OXI     DOREDUCTASE
```

```
32 1aoz_A 106   400   3   0   1 HHH        Cu 344   56   -1   -1   -1   -1   -1 eee    4 1aoz   0 0.1 1.9   ...       |   U     1.10.3.3)        OXI         DOREDUCTASE(O
33 1yai_C   45    80   4   1   0 HHHH       Cu    2   23   55   -1   -1   -1   -1 deee   5 1yai   0 0.4 1.9   ....     |    Z    1.15.1.1;        OXI       DOREDUCTASE
34 1a65_A 111   340   4   0   1 HHHH       Cu 288    2   50   -1   -1   -1   -1 eeee   5 1a65   0 0.2 2.2   ....     |    U     1.10.3.2;       OXI            DOREDUCTASE
35 1phm_   242    72   3   1   0 HHM        Cu    2   70   -1   -1   -1   -1   -1 ee.    4 1phm   0 0.3 1.9   ...       |   Z     1.14.17.3;       MON         OOXYGENASE
36 2trx_A    1     1   3   1   0 ##D        Cu    1    0   -1   -1   -1   -1   -1 ...    4 2trx   0 0.4 1.7   ...       |   Z                    CTRON  TRANSPO ELE
37 1gof_   272   309   3   0   1 YHH        Cu 224   85   -1   -1   -1   -1   -1 .ee    4 1gof   1 0.2 1.7   ...       |   U     1.1.3.9) (       OXI      DOREDUCTASE(O
```

```
no   cngpname     nsp  np  nw  nn  dons     met  sd1  sd2   sd3  sd4  sd5  sd6  sd7  metal    startaa    |  conf_efimov

   1  1oac_A  524   165   3   0   0   HHH       Cu   -1    2  163   -1   -1   -1   -1  CU   801  HISA  524  bbbbb

   2  1ehk_B  149     8   3   0   1  CCH        Cu    4    4   -1   -1   -1   -1   -1  CU1  802  CYSB  149  bbbkbbkjakakk

   3  1cyx_   207     8   4   0   1  COCH       Cu    2    2    4   -1   -1   -1   -1  CU1  201  CYS   207  bbbkbbkjadaak

   4  2cua_A  149     8   4   0   1  COCH       Cu    2    2    4   -1   -1   -1   -1  CU2  169  CYSA  149  bbbkbbkjakaak

   5  2occ_B  196     8   4   0   1  COCH       Cu    2    2    4   -1   -1   -1   -1  CU   229  CYSB  196  bbbd.bkjadakd

   6  1cyx_   172    39   3   0   1  HCC        Cu   35    4   -1   -1   -1   -1   -1  CU2  201  HIS   172  bbbbb-bbbkbbkja

   7  1ehk_B  114    46   4   0   1  HCCM       Cu   35    4    7   -1   -1   -1   -1  CU2  802  HISB  114  bbbbb-bbbkbbkja-kkbbb

   8  2cua_A  114    46   4   0   1  HCCM       Cu   35    4    7   -1   -1   -1   -1  CU1  169  HISA  114  bbbbb-bbbkbbkja-akbbb

   9  2occ_B  161    46   4   0   1  HCCM       Cu   35    4    7   -1   -1   -1   -1  CU   228  HISB  161  bbbbb-bbbd.bkja-kddbb

  10  2mta_A   53    42   3   0   0  HCH        Cu   39    3   -1   -1   -1   -1   -1  CU     0  HISA   53  bbbbb-bbbakdkk

  11  1a65_A  396    61   3   0   0  HCH        Cu   56    5   -1   -1   -1   -1   -1  CU     1  HISA  396  bbbbb-b.bkbakaaa

  12  1pmy_    40    46   4   0   0  HCHM       Cu   38    3    5   -1   -1   -1   -1  CU   124  HIS    40  bgbbb-bbbaakaakgbbb

  13  1f56_A   34    50   4   0   0  HCHM       Cu   40    5    5   -1   -1   -1   -1  CU    92  HISA   34  gkbbb-bbbkbbgaaakgd.b

  14  1plc_    37    55   4   0   0  HCHM       Cu   47    3    5   -1   -1   -1   -1  CU   100  HIS    37  bbbbb-bbbakkaakgbb.

  15  1aoz_A  445    72   4   0   0  HCHM       Cu   62    5    5   -1   -1   -1   -1  CU   701  HISA  445  bbbbb-bbbkbaaaaakgd.b

  16  1qhq_A   57    75   4   0   0  HCHM       Cu   65    5    5   -1   -1   -1   -1  CU   141  HISA   57  bbbbb-bbbkbbgaaakgbb.

  17  1jer_    46    53   4   0   0  HCHQ       Cu   43    5    5   -1   -1   -1   -1  CU   110  HIS    46  gabbb-bbbkb.kaaakgdbb

  18  1e30_A   85    63   4   0   0  HCHQ       Cu   53    5    5   -1   -1   -1   -1  CU   156  HISA   85  bbbbb-bbbkbbgaaakgb.b

  19  1bxa_A   53    45   3   0   0  HCM        Cu   39    6   -1   -1   -1   -1   -1  CU   107  HISA   53  bbbbb-bbbakdkkbbb

  20  1aoz_A   60   388   2   0   1  HH         Cu  388   -1   -1   -1   -1   -1   -1  CU4  703  HISA   60  bbbbb-bbbbb

  21  1phm_   107    65   3   0   0  HHH        Cu    1   64   -1   -1   -1   -1   -1  CU   357  HIS   107  ababbb-bbbba

  22  1yai_B   45    80   3   0   0  HHH        Cu    2   78   -1   -1   -1   -1   -1  CU   152  HISB   45  bbbbbab-bbbab

  23  1ksi_A  442   161   3   2   0  HHH        Cu    2  159   -1   -1   -1   -1   -1  CU   650  HISA  442  bbbbbbb-bbbbb

  24  1lla_   173    31   3   0   0  HHH        Cu    4   27   -1   -1   -1   -1   -1  CU   629  HIS   173  aaaaaakaa-aaaaa

  25  1bt3_A  240    34   3   0   2  HHH        Cu    4   30   -1   -1   -1   -1   -1  CU3  500  HISA  240  aaaaaaaaa-aaaaa

  26  1lla_   324    40   3   0   0  HHH        Cu    4   36   -1   -1   -1   -1   -1  CU   630  HIS   324  daaaaaaaa-aaaaa

  27  1bt3_A   88    30   3   0   2  HHH        Cu   21    9   -1   -1   -1   -1   -1  CU2  500  HISA   88  aaaaakkdgkbbbakdkabbbbabg.kkaaaaaaa

  28  1aoz_A   62   446   3   0   1  HHH        Cu   42  404   -1   -1   -1   -1   -1  CU3  702  HISA   62  bbbgb-.b.bd-bbkba

  29  1a65_A   66   387   3   0   0  HHH        Cu   43  344   -1   -1   -1   -1   -1  CU     3  HISA   66  bbbgb-.b.bk-.bkba

  30  1ehk_A  233    50   3   0   1  HHH        Cu   49    1   -1   -1   -1   -1   -1  CU   803  HISA  233  aaaaa-akkkka

  31  2occ_A  240    51   3   0   1  HHH        Cu   50    1   -1   -1   -1   -1   -1  CU   517  HISA  240  aaaaa-akkkdk
```

```
32 1aoz_A 106   400   3   0   1 HHH        Cu 344   56   -1   -1   -1   -1   -1 CU2 702 HISA 106 .bdbj-bbbgb-bbbbk
33 1yai_C   45    80   4   1   0 HHHH       Cu    2   23   55   -1   -1   -1   -1 CU   152 HISC   45 bbbbbab-jbbbd-bbbab
34 1a65_A 111   340   4   0   1 HHHH       Cu 288    2   50   -1   -1   -1   -1 CU     2 HISA 111 .bkbj-bbbbbgb-bb.bk
35 1phm_  242    72   3   1   0 HHM        Cu    2   70   -1   -1   -1   -1   -1 CU   358 HIS   242 bbbb.ak-bbbbb
36 2trx_A    1     1   3   1   0 ##D        Cu    1    0   -1   -1   -1   -1   -1 CU   109 SERA    1    .kkb
37 1gof_  272   309   3   0   1 YHH        Cu 224   85   -1   -1   -1   -1   -1 CU   700 TYR   272 bbgbb-.bgbb-b.ggb
```

```
no   cngpname     nsp  np  nw  nn  dons     met  sd1  sd2  sd3  sd4  sd5  sd6  sd7  his   cn  pdb  cn2  rms  res   carbi    othdon   ecno

   1 1j79_A 139    38   2   1   0   HH        Zn   -1   38   -1   -1   -1   -1   -1 .de    4 1j79   1 0.2 1.7   ...       |   Z      3.5.2.3;          HYD  ROLASE

   2 1ppt_     1      0   2   0   0 #O         Zn    0   -1   -1   -1   -1   -1   -1 ..     2 1ppt   0 0.2 1.4   ..        |                 CREATIC HORMOPAN

   3 1hw7_A 232      2   2   0   0 CC         Zn    2   -1   -1   -1   -1   -1   -1 ..     2 1hw7   0 0.2 2.2   ..        |                          PERONECHA

   4 1h7n_A 133    10   3   1   0 CCC        Zn    2    8   -1   -1   -1   -1   -1 ...    4 1h7n   0 0.1 1.6   ...       |   Z      4.2.1.24;        DEH  YDRATASE

   5 1a1r_A 123    48   3   1   0 CCC        Zn    2   46   -1   -1   -1   -1   -1 ...    4 1a1r   0 0.2 2.5   ...       |   Z                      AL PROTEIN

   6 1i3j_A 151    16   4   0   0 CCCC       Zn    2   11    3   -1   -1   -1   -1 ....   4 1i3j   0 0.1 2.2   ....      |          3.1.-.-;          HYROLASE/DNA

   7 2occ_F   60    25   4   0   0 CCCC       Zn    2   20    3   -1   -1   -1   -1 ....   4 2occ   0 0.1 2.3   ....      |          1.9.3.1;          OXIDOREDUCTASE

   8 1e3j_A   96    14   4   0   0 CCCC       Zn    3    3    8   -1   -1   -1   -1 ....   4 1e3j   0 0.2 2.3   ....      |                          DOREDUOXIASE

   9 1het_A   97    14   4   0   0 CCCC       Zn    3    3    8   -1   -1   -1   -1 ....   4 1het   0 0.2 1.1   ....      |          1.1.1.1;          OXIREDUCTASE

  10 1zme_C   34    16   4   0   1 CCCC       Zn    3    7    6   -1   -1   -1   -1 ....   5 1zme   0 0.4 2.5   ....      |    J                        COLEX (TRANSCR

  11 1hwt_C   64    17   4   0   1 CCCC       Zn    3    7    7   -1   -1   -1   -1 ....   5 1hwt   0 0.3 2.5   ....      |    J                       PLOEX (ACTIVAT

  12 1f4l_A 145    16   4   0   0 CCCC       Zn    3   10    3   -1   -1   -1   -1 ....   4 1f4l   0 0.3 1.9   ....      |          6.1.1.10;        HYD  ROLASE

  13 1dsz_A1135    20   4   0   0 CCCC       Zn    3   14    3   -1   -1   -1   -1 ....   4 1dsz   0 0.1 1.7   ....      |                             TNSCRIPTION/DN

  14 1dcq_A 264    23   4   0   0 CCCC       Zn    3   17    3   -1   -1   -1   -1 ....   4 1dcq   0 0.1 2.1   ....      |                             MET BINDING PR

  15 1ee8_A 238    23   4   0   0 CCCC       Zn    3   17    3   -1   -1   -1   -1 ....   4 1ee8   0 0.1 1.9   ....      |                             DNBINDING PROT

  16 1rmd_    26    23   4   0   0 CCCC       Zn    3   17    3   -1   -1   -1   -1 ....   4 1rmd   0 0.1 2.1   ....      |                            -BDINDING PROT

  17 1zin_   130    23   4   0   0 CCCC       Zn    3   17    3   -1   -1   -1   -1 ....   4 1zin   0 0.1 1.6   ....      |          2.7.4.3;          PHOSPHOTRANSFERA

  18 1vfy_A 192    33   4   0   0 CCCC       Zn    3   27    3   -1   -1   -1   -1 ....   4 1vfy   0 0.1 1.1   ....      |                            NSTRART PROTEI

  19 1e7l_A   23    38   4   0   0 CCCC       Zn    3   32    3   -1   -1   -1   -1 ....   4 1e7l   0 0.1 1.3   ....      |          3.1.-.-;          ENDONUCLEASE

  20 1ile_   461    43   4   1   0 CCCC       Zn    3   38    2   -1   -1   -1   -1 ....   5 1ile   0 0.8 2.5   ....      |   Z                       NOACAMIL-TRNA S

  21 1hxr_A   23    74   4   0   0 CCCC       Zn    3   68    3   -1   -1   -1   -1 ....   4 1hxr   0 0.1 1.6   ....      |                            ALMEBINDING PR

  22 1ile_   181   211   4   1   0 CCCC       Zn    3  205    3   -1   -1   -1   -1 ....   5 1ile   0 0.7 2.5   ....      |   Z                        NOAMCYL-TRNA S

  23 4mt2_    15    14   4   0   0 CCCC       Zn    4    5    5   -1   -1   -1   -1 ....   4 4mt2   0 0.2 2.0   ....      |                           ALLOMETHIONEIN

  24 1j8f_A 195    29   4   0   0 CCCC       Zn    5   21    3   -1   -1   -1   -1 ....   4 1j8f   0 0.1 1.7   ....      |                           E  REGULATION,

  25 1qf8_A 109    31   4   0   0 CCCC       Zn    5   23    3   -1   -1   -1   -1 ....   4 1qf8   0 0.1 1.7   ....      |          2.7.1.37;        TRA   NSFERASE

  26 1d09_B 109    32   4   0   0 CCCC       Zn    5   24    3   -1   -1   -1   -1 ....   4 1d09   0 0.1 2.1   ....      |          2.1.3.2;          TRA   NSFERASE

  27 1hc7_A 427    34   4   0   0 CCCC       Zn    5   26    3   -1   -1   -1   -1 ....   4 1hc7   0 0.1 2.4   ....      |          6.1.1.15;        AMI    NOACYL-TRNA S

  28 4mt2_     7    19   4   0   0 CCCC       Zn    6    2   11   -1   -1   -1   -1 ....   4 4mt2   0 0.1 2.0   ....      |                           ALLOMETHIONEIN

  29 1dsz_A1171    19   4   0   0 CCCC       Zn    6   10    3   -1   -1   -1   -1 ....   4 1dsz   0 0.1 1.7   ....      |                             TNSCRIPTION/DN

  30 1a5t_    50    15   4   0   0 CCCC       Zn    9    3    3   -1   -1   -1   -1 ....   4 1a5t   0 0.1 2.2   ....      |          2.7.7.7;          CZIFINGER

  31 1zme_C   34    26   4   0   1 CCCC       Zn   16    3    7   -1   -1   -1   -1 ....   5 1zme   0 0.4 2.5   ....      |    J                        COPLEX (TRANSCR
```

```
32 1hwt_C   64    29   4   0   1 CCCC       Zn   17    3    9  -1   -1   -1   -1 ....  5 1hwt  0 0.3 2.5   ....     |   J                        COMPLEX (ACTIVAT
33 1pud_   318    31   4   0   0 CCCH       Zn    2    3   26  -1   -1   -1   -1 ...d  4 1pud  0 0.1 1.9   ....     |         2.4.2.29;        TRA   NSFERASE
34 2hrv_A   52    62   4   0   0 CCCH       Zn    2   58    2  -1   -1   -1   -1 ...d  4 2hrv  0 0.1 2.0   ....     |         3.4.22.29;       HYD     ROLASE
35 1a8h_   127    20   4   0   0 CCCH       Zn    3   14    3  -1   -1   -1   -1 ...d  4 1a8h  0 0.2 2.0   ....     |         6.1.1.10;        AMI   NOACYL-TRNA S
36 1vfy_A 176    27   4   0   0 CCCH       Zn    3   21    3  -1   -1   -1   -1 ...d  4 1vfy  0 0.1 1.1   ....     |                           TRANSPORT PROTEI
37 1a73_A   41    69   4   0   0 CCCH       Zn   59    5    5  -1   -1   -1   -1 ...d  4 1a73  0 0.1 1.8   ....     |                           COMPLEX (HOMING
38 1fn9_A   51    22   4   0   0 CCHC       Zn    3   17    2  -1   -1   -1   -1 ..e.  4 1fn9  0 0.2 1.8   ....     |                         VIRUS/VIRAL PROT
39 1g73_C 300    27   4   0   0 CCHC       Zn    3   17    7  -1   -1   -1   -1 ..e.  4 1g73  0 0.2 2.0   ....     |                           APOPTOSIS/APOPTO
40 1ptq_   244    28   4   0   0 CCHC       Zn    3   22    3  -1   -1   -1   -1 ..d.  4 1ptq  0 0.3 2.0   ....     |         2.7.1.-;         PHOSPHOTRANSFERA
41 1a73_A 125    13   4   0   0 CCHC       Zn    7    2    4  -1   -1   -1   -1 ..d.  4 1a73  0 0.2 1.8   ....     |                           COMPLEX (HOMING
42 1a1i_A 137    20   4   0   0 CCHH       Zn    3   13    4  -1   -1   -1   -1 ..ee  4 1a1i  0 0.1 1.6   ....     |                           COMPLEX (ZINC FI
43 1a1i_A 165    20   4   0   0 CCHH       Zn    3   13    4  -1   -1   -1   -1 ..ee  4 1a1i  0 0.1 1.6   ....     |                           COMPLEX (ZINC FI
44 1guq_A   52   112   4   0   0 CCHH       Zn    3   60   49  -1   -1   -1   -1 ..dd  4 1guq  0 0.1 1.8   ....     |         2.7.7.10;        NUC        LEOTIDYLTRANS
45 1rmd_     91    21   4   0   0 CCHH       Zn    5   12    4  -1   -1   -1   -1 ..ee  4 1rmd  0 0.1 2.1   ....     |                           DNA-BINDING PROT
46 1a1i_A 107    22   4   0   0 CCHH       Zn    5   13    4  -1   -1   -1   -1 ..ee  4 1a1i  0 0.1 1.6   ....     |                           COMPLEX (ZINC FI
47 1clc_   155    42   4   0   0 CCHH       Zn   18    1   23  -1   -1   -1   -1 ..de  4 1clc  0 0.3 1.9   ....     |         3.2.1.4;         GLY  COSYL HYDROLA
48 1g5c_A   32    58   3   1   0 CHC        Zn   55    3   -1  -1   -1   -1   -1 .e.   4 1g5c  0 0.1 2.1   ...      |  Z      4.2.1.1;         LYA  SE
49 1ekj_D 160    63   3   0   0 CHC        Zn   60    3   -1  -1   -1   -1   -1 .e.   3 1ekj  0 0.1 1.9   ...      |         4.2.1.1          LYASE
50 1g71_A 106    11   4   0   0 CHCC       Zn    2    6    3  -1   -1   -1   -1 .e..  4 1g71  0 0.2 2.3   ....     |                          REPLICATION
51 1rmd_     41    23   4   0   0 CHCC       Zn    2   18    3  -1   -1   -1   -1 .d..  4 1rmd  0 0.1 2.1   ....     |                           DNA-BINDING PROT
52 1d0q_A   40    24   4   0   0 CHCC       Zn    3   18    3  -1   -1   -1   -1 .d..  4 1d0q  0 0.2 1.7   ....     |         2.7.7.-;         TRA  NSFERASE
53 1tup_A 176    66   4   0   0 CHCC       Zn    3   59    4  -1   -1   -1   -1 .d..  4 1tup  0 0.1 2.2   ....     |                           COMPLEX (TUMOR S
54 1vsr_A   66    51   4   0   0 CHCC       Zn    5    2   44  -1   -1   -1   -1 .d..  4 1vsr  0 0.2 1.8   ....     |         3.1.-.-;         HYD ROLASE
55 1rmd_      2    29   4   0   0 CHCH       Zn    4   23    2  -1   -1   -1   -1 .d.d  4 1rmd  0 0.1 2.1   ....     |                           DNA-BINDING PROT
56 1kev_A   37   113   3   0   0 CHD        Zn   22   91   -1  -1   -1   -1   -1 .e.   3 1kev  1 0.1 2.0   ...      |         1.1.1.2;         OXI  DOREDUCTASE
57 1e3j_A   41    26   3   1   0 CHE        Zn   25    1   -1  -1   -1   -1   -1 .e.   4 1e3j  0 0.2 2.3   ...      |  Z                       OXIDOREDUCTASE
58 1kev_B   37   113   4   0   0 CHED       Zn   22    1   90  -1   -1   -1   -1 .e..  4 1kev  0 0.3 2.0   ....     |         1.1.1.2;         OXI   DOREDUCTASE
59 1bs4_C1090    46   3   0   1 CHH        Zn   42    4   -1  -1   -1   -1   -1 .ee   4 1bs4  0 0.1 1.9   ...      |  U      3.5.1.31;        HYD      ROLASE
60 1evl_A 334   177   3   0   2 CHH        Zn   51  126   -1  -1   -1   -1   -1 .ed   5 1evl  0 0.3 1.5   ...      |  BU    6.1.1.3;          LIG      ASE
61 1czf_B 308    28   2   1   0 DD         Zn   28   -1   -1  -1   -1   -1   -1 ..    3 1czf  1 0.1 1.7   ..       | Z       3.2.1.15         HYD ROLASE
62 1frp_A 118   162   3   1   1 DDE        Zn    3  159   -1  -1   -1   -1   -1 ...   5 1frp  0 0.4 2.0   ...      |  UZ    3.1.3.11)         HYD      ROLASE(PHOSPH
63 1ete_B   40     2   2   1   0 DE         Zn    2   -1   -1  -1   -1   -1   -1 ..    3 1ete  1 0.3 2.2   ..       | Z                    OKINE  CYT
64 1cy5_A   27    13   2   1   0 DE         Zn   13   -1   -1  -1   -1   -1   -1 ..    3 1cy5  0 0.3 1.3   .b       | Z                        APOPTOSIS
```

```
 65  1ete_C   14    64   2   0   0  DE          Zn   64   -1   -1   -1   -1   -1   -1 ..      2  1ete   0  0.2  2.2   ..        |                         OKINE  CYT
 66  1amp_   117   139   3   1   0  DEH         Zn   35  104   -1   -1   -1   -1   -1 ..e    4  1amp   0  0.3  1.8   .b.       |  Z     3.4.11.10)      HYD          ROLASE(AMINOP
 67  1cg2_A  141   244   3   1   0  DEH         Zn   35  209   -1   -1   -1   -1   -1 ..e    4  1cg2   0  0.2  2.5   .b.       |  Z     3.4.17.11;

METALLOCARBOXYPE

 68  1qh5_B   58   115   4   1   0  DHDH        Zn    1   75   39   -1   -1   -1   -1 .e.e   5  1qh5   0  0.2  1.5   ....      |   Z    3.1.2.6;        HYD        ROLASE
 69  1ush_     41   213   4   0   1  DHDQ        Zn    2   41  170   -1   -1   -1   -1 .e..   5  1ush   0  0.2  1.7   ....      |   U    3.1.3.5;        HYD      ROLASE
 70  1sml_A   88   137   3   2   0  DHH         Zn    1  136   -1   -1   -1   -1   -1 .ee    5  1sml   0  0.2  1.7   ...       |  ZZ    3.5.2.6;        HYD       ROLASE
 71  1ed8_A  327    85   3   0   2  DHH         Zn    4   81   -1   -1   -1   -1   -1 .ee    5  1ed8   0  0.4  1.8   b..       |  UU    3.1.3.1;        HYD       ROLASE
 72  1ew2_A  316   116   3   0   2  DHH         Zn    4  112   -1   -1   -1   -1   -1 .ee    5  1ew2   0  0.4  1.8   b..       |  UU    3.1.3.1         HYD           ROLASE
 73  1ah7_     55    67   4   1   0  DHHD        Zn   14   49    4   -1   -1   -1   -1 .de.   5  1ah7   0  0.2  1.5   ....      |   Z    3.1.4.3;        HYD     ROLASE
 74  1ak0_     45    75   4   0   0  DHHD        Zn   15   56    4   -1   -1   -1   -1 .de.   4  1ak0   0  0.5  1.8   ....      |        3.1.30.1        END    ONUCLEASE
 75  1ush_     84   168   4   0   2  DNHH        Zn   32  101   35   -1   -1   -1   -1 ..ed   6  1ush   0  0.3  1.7   ....     |   UU  3.1.3.5;        HYD          ROLASE
 76  1byf_A   52    74   2   0   2  DO          Zn   74   -1   -1   -1   -1   -1   -1 ..      4  1byf   1  0.4  2.0   ..        | UU                          SUGAR BINDING  PR
 77  1lam_   255    79   4   1   1  DODE        Zn   77    0    2   -1   -1   -1   -1 ....   6  1lam   0  0.4  1.6   ....      |   JZ  3.4.11.1        HYD     ROLASE  (ALPHA
 78  1ew2_A   42   315   3   0   1  DSD         Zn   50  265   -1   -1   -1   -1   -1 ...    4  1ew2   1  0.4  1.8   b..       |  U     3.1.3.1         HYD       ROLASE
 79  1taf_B   31     4   2   1   0  ED          Zn    4   -1   -1   -1   -1   -1   -1 ..      3  1taf   0  0.4  2.0   .b        | Z                     PLEX (TWOCOM TRA
 80  1noy_B  114   210   2   0   1  ED          Zn  210   -1   -1   -1   -1   -1   -1 ..      3  1noy   0  0.6  2.2   ..        | U      2.7.7.7;        COM     PLEX  (NUCLEOT
 81  1d8w_B  234   100   4   0   0  EDHD        Zn   33   27   40   -1   -1   -1   -1 ..d.   4  1d8w   0  0.1  1.6   ....     |        5.3.1.14;       ISO         MERASE
 82  1qtw_A  145   116   4   1   0  EDHE        Zn   34   37   45   -1   -1   -1   -1 ..d.   5  1qtw   0  0.2  1.0   ....     |   Z    3.1.21.2;       HYD           ROLASE
 83  3cao_A   26     3   2   0   0  EE          Zn    3   -1   -1   -1   -1   -1   -1 ..      2  3cao   1  0.4  1.6   b.        |                         CTRON  TRANSPO
 84  1ec5_A   10    29   3   0   0  EEH         Zn   26    3   -1   -1   -1   -1   -1 ..d    3  1ec5   0  0.2  2.5   b..       |                           NOVODEPROTEIN
 85  1ete_C   58    22   2   1   0  EH          Zn   22   -1   -1   -1   -1   -1   -1 .e      3  1ete   0  0.6  2.2   ..        | Z                        OKINECYT
 86  1cvr_A  152    59   2   1   0  EH          Zn   59   -1   -1   -1   -1   -1   -1 .e      3  1cvr   1  0.5  2.0   ..        | Z      3.4.22.37;      HYD   ROLASE
 87  1bf6_A  125    61   3   0   0  EHH         Zn   33   28   -1   -1   -1   -1   -1 .de    3  1bf6   0  0.2  1.7   ...       |                              PHOSPHOTRIESTERA
 88  1dqs_B  194    93   3   0   2  EHH         Zn   77   16   -1   -1   -1   -1   -1 .ee    5  1dqs   0  0.1  1.8   ...       |  UU    4.6.1.3;        LYA       SE
 89  1e4c_P   73    82   4   1   0  EHHH        Zn   19    2   61   -1   -1   -1   -1 .eee   5  1e4c   0  0.2  1.7   ....      |   Z    4.1.2.17;       ALD       OLASE  (CLASS
 90  1a2p_C   60     2   2   1   0  EK          Zn    2   -1   -1   -1   -1   -1   -1 ..      3  1a2p   0  0.2  1.5   b.        | Z      3.1.27.-;        RIBONUCLEASE
 91  1ctt_   102    30   3   1   0  HCC         Zn   27    3   -1   -1   -1   -1   -1 d..     4  1ctt   0  0.2  2.2   ...       |  Z     3.5.4.5) C      HYD ROLASE
 92  1btk_A  143    22   4   0   0  HCCC        Zn   11    1   10   -1   -1   -1   -1 d...   4  1btk   0  0.1  1.6   ....     |        2.7.1.112;      TRA      NSFERASE
 93  1gpc_     64    26   4   0   0  HCCC        Zn   13   10    3   -1   -1   -1   -1 e...   4  1gpc   0  0.4  2.2   ....      |                             DNA BINDING  PROT
 94  1ptq_   231    49   4   0   0  HCCC        Zn   30    3   16   -1   -1   -1   -1 d...   4  1ptq   0  0.3  2.0   ....      |        2.7.1.-;        PHO  SPHOTRANSFERA
 95  1amp_     97    82   3   1   0  HDD         Zn   20   62   -1   -1   -1   -1   -1 e..     4  1amp   0  0.2  1.8   ..b       |  Z     3.4.11.10)      HYD       ROLASE(AMINOP
 96  1cg2_A  112    88   3   1   0  HDE         Zn   29   59   -1   -1   -1   -1   -1 e..    4  1cg2   0  0.2  2.5   ..b       |  Z     3.4.17.11;      MET       ALLOCARBOXYPE
```

```
 97 1j9y_A   79    42   3   1   0 HDE        Zn   32   10   -1   -1   -1   -1   -1 e..    4 1j9y   0 0.2 1.9   ..b       |   Z     3.2.1.78;        HYD     ROLASE
 98 1i76_A  147    15   3   0   0 HDH        Zn    2   13   -1   -1   -1   -1   -1 e.e    3 1i76   0 0.1 1.2   ...       |         3.4.24.34;      HYD     ROLASE
 99 1vhh_   141    42   3   1   0 HDH        Zn    7   35   -1   -1   -1   -1   -1 e.d    4 1vhh   1 0.1 1.7   ...       |   Z                          SIGNALING PROTE
100 1lbu_   154    43   3   1   0 HDH        Zn    7   36   -1   -1   -1   -1   -1 e.d    4 1lbu   0 0.2 1.8   ...       |   Z     3.4.17.8         HYD    ROLASE
101 1qtw_A  182    49   3   2   0 HDH        Zn   47    2   -1   -1   -1   -1   -1 e.e    5 1qtw   0 0.2 1.0   .b.       |  ZZ    3.1.21.2;        HYD           ROLASE
102 2cua_B  117     2   2   0   0 HE         Zn    2   -1   -1   -1   -1   -1   -1 e.     2 2cua   1 0.5 1.6   ..        |         1.9.3.1;         ELECTRON TRANSPO
103 1i1i_P  160     4   2   1   0 HE         Zn    4   -1   -1   -1   -1   -1   -1 e.     3 1i1i   0 0.2 2.3   ..        | Z       3.4.24.16;      HYDROLASE
104 1ak0_    15   166   2   1   0 HE         Zn  166   -1   -1   -1   -1   -1   -1 e.     3 1ak0   0 0.3 1.8   .b        | Z       3.1.30.1         END    ONUCLEASE
105 2a0b_   728    30   3   0   0 HEE        Zn   26    4   -1   -1   -1   -1   -1 e..    3 2a0b   1 0.1 1.6   ...       |         2.7.3.-;         SENSORY TRANSDUC
106 1pyt_B   69   127   3   1   0 HEH        Zn    3  124   -1   -1   -1   -1   -1 d.d    4 1pyt   0 0.4 2.3   .b.       |  Z     3.4.17.1;        TER        NARY COMPLEX
107 2ctc_    69   127   3   1   0 HEH        Zn    3  124   -1   -1   -1   -1   -1 d.d    4 2ctc   0 0.2 1.4   .b.       |  Z     3.4.17.1)        HYD      ROLASE(C-TERM
108 1dq3_A  156     2   2   0   0 HH         Zn    2   -1   -1   -1   -1   -1   -1 de     2 1dq3   0 0.2 2.1   ..        |                          HYDROLASE
109 1fio_A   75     3   2   0   0 HH         Zn    3   -1   -1   -1   -1   -1   -1 dd     2 1fio   0 0.2 2.1   ..        |                         MEMBRANE PROTEIN
110 1hzy_A  201    29   2   2   1 HH         Zn   29   -1   -1   -1   -1   -1   -1 de     5 1hzy   0 0.2 1.3   ..        | UZZ    3.1.8.1;         HYD         ROLASE
111 1j79_A   16   234   3   1   1 HH  D      Zn    2   -1  232   -1   -1   -1   -1 ee .   5 1j79   0 0.2 1.7   .. .      | U Z    3.5.2.3;         HYD         ROLASE
112 1j98_A   54    72   3   1   0 HHC        Zn    4   68   -1   -1   -1   -1   -1 ee.    4 1j98   0 0.1 1.2   ...       |   Z                         SIGNALING PROTEI
113 1lba_    17   113   3   1   0 HHC        Zn  105    8   -1   -1   -1   -1   -1 dd.    4 1lba   0 0.6 2.2   ...       |  Z     3.5.1.28)        HYD      ROLASE(ACTING
114 1ia9_A1751    63   4   0   0 HHCC       Zn   57    2    4   -1   -1   -1   -1 de..   4 1ia9   0 0.2 2.0   ....      |         2.7.1.37;        TRA       NSFERASE
115 1hzy_B   55   246   3   1   1 HHD        Zn    2  244   -1   -1   -1   -1   -1 ee.    5 1hzy   0 0.1 1.3   ...       |  UZ    3.1.8.1;         HYD          ROLASE
116 1c7k_A   83    10   3   1   0 HHD        Zn    4    6   -1   -1   -1   -1   -1 ee.    4 1c7k   0 0.1 1.0   ...       |   Z     3.4.24.-          HYD    ROLASE
117 1g12_A  117    13   3   2   0 HHD        Zn    4    9   -1   -1   -1   -1   -1 ee.    5 1g12   0 0.2 1.6   ..b       |  ZZ    3.4.24.20        HYD          ROLASE
118 1ak0_   126    27   3   0   1 HHD        Zn   23    4   -1   -1   -1   -1   -1 ee.    4 1ak0   0 0.4 1.8   ..b       |   U     3.1.30.1         END       ONUCLEASE
119 1eu3_A  162    42   3   0   0 HHD        Zn   40    2   -1   -1   -1   -1   -1 de.    3 1eu3   1 0.1 1.7   ...       |                           IMMUNE SYSTEM
120 1f0j_A  238   154   4   1   0 HHDD       Zn   36    1  117   -1   -1   -1   -1 ee..   5 1f0j   0 0.2 1.8   ....      |    Z    3.1.4.17;        HYD           ROLASE
121 1hs6_A  295    23   3   0   2 HHE        Zn    4   19   -1   -1   -1   -1   -1 ee.    5 1hs6   1 0.2 2.0   ...       |  UU    3.3.2.6;         HYD        ROLASE
122 1ezm_   140    24   3   1   0 HHE        Zn    4   20   -1   -1   -1   -1   -1 ee.    4 1ezm   1 0.2 1.5   ...       |  Z     3.4.24.26)       HYD        ROLASE
123 1i1i_P  474    29   3   1   0 HHE        Zn    4   25   -1   -1   -1   -1   -1 ee.    4 1i1i   0 0.4 2.3   ..b       |  Z     3.4.24.16;       HYD      ROLASE
124 1dmt_A  583    63   3   0   1 HHE        Zn    4   59   -1   -1   -1   -1   -1 ee.    4 1dmt   0 0.4 2.1   ..b       |  U     3.4.24.11;       HYD           ROLASE
125 1hr6_B   70    80   3   1   0 HHE        Zn    4   76   -1   -1   -1   -1   -1 ee.    4 1hr6   0 0.1 2.5   ..b       |  Z     3.4.24.64;       HYD       ROLASE
126 1qtw_A   69    76   3   1   0 HHE        Zn   40   36   -1   -1   -1   -1   -1 ee.    4 1qtw   0 0.1 1.0   ...       |  Z     3.1.21.2;        HYD       ROLASE
127 1bf6_A   12   231   4   0   0 HHED       Zn    2  111  118   -1   -1   -1   -1 ee..   4 1bf6   0 0.3 1.7   ....      |                          PHO  SPHOTRIESTERA
128 1psz_A   67   213   4   0   0 HHED       Zn   72   66   75   -1   -1   -1   -1 ee..   4 1psz   1 0.3 2.0   ..b.      |                          IMM  UNE SYSTEM
129 1koq_A   92    19   3   1   0 HHH        Zn    2   17   -1   -1   -1   -1   -1 eed    4 1koq   0 0.2 1.9   ...       |  Z     4.2.1.1;         LYA      SE
```

```
130 2bc2_A   86    63   3   2   0 HHH        Zn    2  61   -1   -1   -1   -1   -1 ede    5 2bc2   0 0.3 1.7   ...       |  ZZ    3.5.2.6;         HYD        ROLASE
131 1sml_A   84    76   3   1   0 HHH        Zn    2  74   -1   -1   -1   -1   -1 ede    4 1sml   1 0.2 1.7   ...       |  Z      3.5.2.6;         HYD      ROLASE
132 1i76_A  197    10   3   0   2 HHH        Zn    4    6   -1   -1   -1   -1   -1 eee    5 1i76   0 0.3 1.2   ...       |  UU    3.4.24.34;       HYD         ROLASE
133 1lml_   264    70   3   1   0 HHH        Zn    4  66   -1   -1   -1   -1   -1 eee    4 1lml  20  0.4 1.9   ...      |  Z     3.4.24.36        LEI      SHMANOLYSIN
134 1b66_B   23    27   3   0   1 HHH        Zn   25    2   -1   -1   -1   -1   -1 eee    4 1b66   0 0.5 1.9   ...       |  U      4.6.1.10;       TET      RAHYDROBIOPTE
135 1ton_    57    42   3   0   0 HHH        Zn   40    2   -1   -1   -1   -1   -1 eee    3 1ton   0 0.1 1.8   ...       |           NUMBER NO       HYD          ROLASE(SERINE
136 1bkc_E  405    10   3   0   2 HHH        Zn    4    6   -1   -1   -1   -1   -1 eee    5 1bkc   0 0.4 2.0   ...       |  UU                        ZN-ENDOPEPTIDASE
137 1qh5_B   54    80   4   1   0 HHHD       Zn    2  54   24   -1   -1   -1   -1 ede.   5 1qh5   0 0.3 1.5   ....     |    Z    3.1.2.6;         HYD         ROLASE
138 1mfm_A   63    20   4   0   0 HHHD       Zn    8    9    3   -1   -1   -1   -1 ddd.   4 1mfm   1 0.1 1.0   ....     |          1.15.1.1;       OXI        DOREDUCTASE
139 1yai_A   70    21   4   0   0 HHHD       Zn    9    9    3   -1   -1   -1   -1 ddd.   4 1yai   1 0.2 1.9   ....     |          1.15.1.1;       OXI    DOREDUCTASE
140 1toa_A   68   211   4   0   0 HHHD       Zn   65   66   80   -1   -1   -1   -1 eee.   4 1toa   0 0.3 1.8   ...b     |                                BIN DING PROTEIN
141 1lam_   250    84   4   1   1 KDDE       Zn    5   18   61   -1   -1   -1   -1 ....   6 1lam   0 0.4 1.6   ....     |    JZ   3.4.11.1         HYD        ROLASE (ALPHA
142 1b0n_A    1    63   3   1   0 #OE        Zn    0   63   -1   -1   -1   -1   -1 ...    4 1b0n   1 0.2 1.9   ...       |   Z                         NSCRIPTION RE
143 1lam_   173    98   2   1   0 OO         Zn   98   -1   -1   -1   -1   -1   -1 ..     3 1lam  22  0.7 1.6   ..        | Z       3.4.11.1         HYD   ROLASE (ALPHA
144 1pmi_   111   174   4   1   0 QHEH       Zn    2   25 147   -1   -1   -1   -1 .e.e   5 1pmi   1 0.1 1.7   ....     |    Z    5.3.1.8          ISO        MERASE
145 1taf_B   45     0   2   2   0 QQ         Zn    0   -1   -1   -1   -1   -1   -1 ..     4 1taf   0 0.5 2.0   ..        | ZZ                       PLEX COM(TWO TRA
146 1j9y_A  208   112   4   0   0 RHDE       Zn    3   72   37   -1   -1   -1   -1 .d..   4 1j9y   0 0.4 1.9   ..b.     |          3.2.1.78;       HYD          ROLASE
147 1ak0_     1   119   4   0   0 #OHD       Zn    0    5 114   -1   -1   -1   -1 ..e.   4 1ak0   0 0.4 1.8   ....     |          3.1.30.1         END    ONUCLEASE
148 1ca1_     1   129   4   1   0 #OHD       Zn    0   10 119   -1   -1   -1   -1 ..e.   5 1ca1   0 0.5 1.9   ....     |    Z    3.1.4.3;         HYD      ROLASE
149 1ah7_     1   121   4   1   0 #OHD       Zn    0   13 108   -1   -1   -1   -1 ..e.   5 1ah7   1 0.2 1.5   ....     |    Z    3.1.4.3;         HYD      ROLASE
150 1fio_A   30     1   3   0   1 ##H        Zn    1    0   -1   -1   -1   -1   -1 . d    3 1fio   0 0.2 2.1   . .       |B                          BRANEMEMPROTEIN
151 1qe3_A    2     1   3   2   1 ##H        Zn    1    0   -1   -1   -1   -1   -1 . d    5 1qe3   0 0.5 1.5   . .       |B ZZ    3.1.1.-;         HYD  ROLASE
```

```
no   cngpname     nsp  np  nw  nn  dons     met  sd1  sd2   sd3  sd4  sd5  sd6  sd7  metal    startaa    |  conf_efimov
   1 1j79_A 139    38   2   1   0   HH        Zn   -1   38   -1   -1   -1   -1   -1 ZN   401 HISA 139 bbbbb
   2 1ppt_     1     0   2   0   0 #O         Zn    0   -1   -1   -1   -1   -1   -1 ZN     1 GLY     1    .bb
   3 1hw7_A 232     2   2   0   0 CC          Zn    2   -1   -1   -1   -1   -1   -1 ZN   240 CYSA 232 bbbk...
   4 1h7n_A 133    10   3   1   0 CCC        Zn    2    8   -1   -1   -1   -1   -1 ZN 1342 CYSA 133 bb.kakkbakgbkbb
   5 1a1r_A 123    48   3   1   0 CCC        Zn    2   46   -1   -1   -1   -1   -1 ZN   901 CYSA 123 bbbkkgb-bbbkk
   6 1i3j_A 151    16   4   0   0 CCCC       Zn    2   11    3   -1   -1   -1   -1 ZN   100 CYSA 151 .bbakgbbbbkkbabbaakbb
   7 2occ_F   60    25   4   0   0 CCCC       Zn    2   20    3   -1   -1   -1   -1 ZN    99 CYSF   60 bbbbabadbkbbbb.bbbb.bbbbbaakgb
   8 1e3j_A   96    14   4   0   0 CCCC       Zn    3    3    8   -1   -1   -1   -1 ZN   901 CYSA   96 bbagkbaaaakgdakkbad
   9 1het_A   97    14   4   0   0 CCCC       Zn    3    3    8   -1   -1   -1   -1 ZN   401 CYSA   97 bbagkbaaakbaabkdbkk
  10 1zme_C   34    16   4   0   1 CCCC       Zn    3    7    6   -1   -1   -1   -1 ZN     1 CYSC   34 bbbaaaaakgbbbkjj.bbaa
  11 1hwt_C   64    17   4   0   1 CCCC       Zn    3    7    7   -1   -1   -1   -1 ZN     1 CYSC   64 kbbaaaaaagbbbkgbbabbaa
  12 1f4l_A 145    16   4   0   0 CCCC       Zn    3   10    3   -1   -1   -1   -1 ZN   701 CYSA 145 .bbaakgbbgbbjkbbkakgb
  13 1dsz_A1135    20   4   0   0 CCCC       Zn    3   14    3   -1   -1   -1   -1 ZN 1121 CYSA1135 ..baakgbbbkjbbggbbbbaaaaa
  14 1dcq_A 264    23   4   0   0 CCCC       Zn    3   17    3   -1   -1   -1   -1 ZN   600 CYSA 264 kbbakkgbbgbkbbbakkgbbbbaaaaa
  15 1ee8_A 238    23   4   0   0 CCCC       Zn    3   17    3   -1   -1   -1   -1 ZN   501 CYSA 238 bbbkakgbbbbbbbbgg.bbb.baaadg
  16 1rmd_    26    23   4   0   0 CCCC       Zn    3   17    3   -1   -1   -1   -1 ZN   118 CYS    26 kbbkakgbbbbgbbbbakgbbbbaaaaa
  17 1zin_   130    23   4   0   0 CCCC       Zn    3   17    3   -1   -1   -1   -1 ZN   219 CYS   130 bbbaakgbbbbaaa.bbabbgbbaakgb
  18 1vfy_A 192    33   4   0   0 CCCC       Zn    3   27    3   -1   -1   -1   -1 ZN   300 CYSA 192 bbbkakgb-bbbaaaaa
  19 1e7l_A   23    38   4   0   0 CCCC       Zn    3   32    3   -1   -1   -1   -1 ZN 1165 CYSA   23 gbbaakgb-bbbaaaaa
  20 1ile_   461    43   4   1   0 CCCC       Zn    3   38    2   -1   -1   -1   -1 ZN 1102 CYS   461 bbba....-bbbak.b
  21 1hxr_A   23    74   4   0   0 CCCC       Zn    3   68    3   -1   -1   -1   -1 ZN   200 CYSA   23 bbbkakgb-bbbaakgb
  22 1ile_   181   211   4   1   0 CCCC       Zn    3  205    3   -1   -1   -1   -1 ZN 1101 CYS   181 bbbaaag.-bbbkabkb
  23 4mt2_    15    14   4   0   0 CCCC       Zn    4    5    5   -1   -1   -1   -1 ZN    67 CYS    15 bdaggkbkbadbkbkkbdb
  24 1j8f_A 195    29   4   0   0 CCCC       Zn    5   21    3   -1   -1   -1   -1 ZN 1001 CYSA 195 bbbkbakagbbbbaaaaaaaaakgbbbbbkakgb
  25 1qf8_A 109    31   4   0   0 CCCC       Zn    5   23    3   -1   -1   -1   -1 ZN   216 CYSA 109 jbbkbakkgg-bbbaa...
  26 1d09_B 109    32   4   0   0 CCCC       Zn    5   24    3   -1   -1   -1   -1 ZN 1313 CYSB 109 bbbkbakbaa-bbbaakgb
  27 1hc7_A 427    34   4   0   0 CCCC       Zn    5   26    3   -1   -1   -1   -1 ZN   490 CYSA 427 bbagbaaaaa-.bbaakgb
  28 4mt2_     7    19   4   0   0 CCCC       Zn    6    2   11   -1   -1   -1   -1 ZN    68 CYS     7 bkbbdkgkbdaggkbkbadbkbkk
  29 1dsz_A1171    19   4   0   0 CCCC       Zn    6   10    3   -1   -1   -1   -1 ZN 1122 CYSA1171 bbbk.kgkbbbbakkakabaaaaa
  30 1a5t_    50    15   4   0   0 CCCC       Zn    9    3    3   -1   -1   -1   -1 ZN   501 CYS    50 akbkdbb.kbbagabaaaaa
  31 1zme_C   34    26   4   0   1 CCCC       Zn   16    3    7   -1   -1   -1   -1 ZN     2 CYSC   34 bbbaaaaakgbbbkjj.bbaaaaakgbbbbb
```

```
32  1hwt_C   64    29   4   0   1  CCCC      Zn   17    3    9   -1   -1   -1   -1  ZN      2  CYSC   64  kbbaaaaaagbbbkgbbabbaaaaakgkakkbbb
33  1pud_   318    31   4   0   0  CCCH      Zn    2    3   26   -1   -1   -1   -1  ZN    400  CYS   318  kkbkbaakka-aaaaa
34  2hrv_A   52    62   4   0   0  CCCH      Zn    2   58    2   -1   -1   -1   -1  ZN    143  CYSA   52  bdbkbkb-bbbaa.b
35  1a8h_   127    20   4   0   0  CCCH      Zn    3   14    3   -1   -1   -1   -1  ZN   1000  CYS   127  bbbaaagbbbbaaabdg.bbabggb
36  1vfy_A  176    27   4   0   0  CCCH      Zn    3   21    3   -1   -1   -1   -1  ZN    301  CYSA  176  kbbkakgbbbbakbbbbbbkakgbbbbaakdk
37  1a73_A   41    69   4   0   0  CCCH      Zn   59    5    5   -1   -1   -1   -1  ZN    202  CYSA   41  bbbbb-bkkgbakbkbkkkbb
38  1fn9_A   51    22   4   0   0  CCHC      Zn    3   17    2   -1   -1   -1   -1  ZN   1001  CYSA   51  bbbkkkgbbk.baakgbkbbbbbbbkg
39  1g73_C  300    27   4   0   0  CCHC      Zn    3   17    7   -1   -1   -1   -1  ZN    501  CYSC  300  bbbaakgbbbbgbbabbbaaaaaaaadakbka
40  1ptq_   244    28   4   0   0  CCHC      Zn    3   22    3   -1   -1   -1   -1  ZN      2  CYS   244  bbbkakgbbbbjabga.bbbaakgbbbbakkak
41  1a73_A  125    13   4   0   0  CCHC      Zn    7    2    4   -1   -1   -1   -1  ZN    201  CYSA  125  akbkjkkkgbkbabbbkb
42  1a1i_A  137    20   4   0   0  CCHH      Zn    3   13    4   -1   -1   -1   -1  ZN    202  CYSA  137  bbbkakgbbbabaaaaaaaakaaag
43  1a1i_A  165    20   4   0   0  CCHH      Zn    3   13    4   -1   -1   -1   -1  ZN    203  CYSA  165  bbbaakgbbbabaaaaaaaakkkb.
44  1guq_A   52   112   4   0   0  CCHH      Zn    3   60   49   -1   -1   -1   -1  ZN    350  CYSA   52  aabkkbbg-kbkkb-bkbb.
45  1rmd_    91    21   4   0   0  CCHH      Zn    5   12    4   -1   -1   -1   -1  ZN    120  CYS    91  bbbkbbgbdbbbbaakaaaaaakbbb
46  1a1i_A  107    22   4   0   0  CCHH      Zn    5   13    4   -1   -1   -1   -1  ZN    201  CYSA  107  bbbkbakbkbbbabaaaaaaaaaaakg
47  1clc_   155    42   4   0   0  CCHH      Zn   18    1   23   -1   -1   -1   -1  ZN    594  CYS   155  kbbgb-bbbgkk-jbba.
48  1g5c_A   32    58   3   1   0  CHC       Zn   55    3   -1   -1   -1   -1   -1  ZN   1001  CYSA   32  bbbkb-bbbbgagk
49  1ekj_D  160    63   3   0   0  CHC       Zn   60    3   -1   -1   -1   -1   -1  ZN   4004  CYSD  160  bbbkb-bbbbgbaa
50  1g71_A  106    11   4   0   0  CHCC      Zn    2    6    3   -1   -1   -1   -1  ZN    339  CYSA  106  babdbbbgbbbaaaaa
51  1rmd_    41    23   4   0   0  CHCC      Zn    2   18    3   -1   -1   -1   -1  ZN    119  CYS    41  bakgbbbbaaaaaaaaaaajdbbaakgb
52  1d0q_A   40    24   4   0   0  CHCC      Zn    3   18    3   -1   -1   -1   -1  ZN    151  CYSA   40  .bbkabbbabbkbbbbaakgbb.bkk.gb
53  1tup_A  176    66   4   0   0  CHCC      Zn    3   59    4   -1   -1   -1   -1  ZN    951  CYSA  176  bbbaaaak-bkbbakbbg
54  1vsr_A   66    51   4   0   0  CHCC      Zn    5    2   44   -1   -1   -1   -1  ZN    201  CYSA   66  bbaaakgbkbkk-baaaa
55  1rmd_     2    29   4   0   0  CHCH      Zn    4   23    2   -1   -1   -1   -1  ZN    117  CYS     2   .kaakbbbkkkbabdbbaaaaaakbbkakgbbb
56  1kev_A   37   113   3   0   0  CHD       Zn   22   91   -1   -1   -1   -1   -1  ZN    353  CYSA   37  bbbaa-b..b.-aa.aa
57  1e3j_A   41    26   3   1   0  CHE       Zn   25    1   -1   -1   -1   -1   -1  ZN    902  CYSA   41  bbbaaaaaaaakg.k.kbbbabbbbb..bbb
58  1kev_B   37   113   4   0   0  CHED      Zn   22    1   90   -1   -1   -1   -1  ZN    353  CYSB   37  bbbaa-b..b.b-aa.aa
59  1bs4_C1090    46   3   0   1  CHH       Zn   42    4   -1   -1   -1   -1   -1  ZN   2003  CYSC1090  bb.ak-aaaaaaakk
60  1evl_A  334   177   3   0   2  CHH       Zn   51  126   -1   -1   -1   -1   -1  ZN      1  CYSA  334  bkaaa-b.bbb-bbbbb
61  1czf_B  308    28   2   1   0  DD        Zn   28   -1   -1   -1   -1   -1   -1  ZN    504  ASPB  308  bbgbbbbgbbjbbbkkbbabbbbbbbgkbbgbb
62  1frp_A  118   162   3   1   1  DDE       Zn    3  159   -1   -1   -1   -1   -1  ZN      5  ASPA  118  bbbabbga-akaaa
63  1ete_B   40     2   2   1   0  DE        Zn    2   -1   -1   -1   -1   -1   -1  ZN   1137  ASPB   40  bbbbaaa
64  1cy5_A   27    13   2   1   0  DE        Zn   13   -1   -1   -1   -1   -1   -1  ZN    102  ASPA   27  aaaaaaaagkbbaaaaaa
```

```
65 1ete_C  14   64  2  0  0 DE        Zn  64  -1  -1  -1  -1  -1  -1 ZN 2136 ASPC  14 bkkaa-aaaka
66 1amp_  117  139  3  1  0 DEH       Zn  35 104  -1  -1  -1  -1  -1 ZN  501 ASP  117 aabda-gakkg-kbgbk
67 1cg2_A 141  244  3  1  0 DEH       Zn  35 209  -1  -1  -1  -1  -1 ZN  500 ASPA 141 aabda-dakdg-gbgbk
68 1qh5_B  58  115  4  1  0 DHDH      Zn   1  75  39  -1  -1  -1  -1 ZN  262 ASPB  58 bakakk-k.kab-babdk
69 1ush_   41  213  4  0  1 DHDQ      Zn   2  41 170  -1  -1  -1  -1 ZN  600 ASP   41 b.gaggd-kjgab-.kgbb
70 1sml_A  88  137  3  2  0 DHH       Zn   1 136  -1  -1  -1  -1  -1 ZN  268 ASPA  88 bakaka-babaa
71 1ed8_A 327   85  3  0  2 DHH       Zn   4  81  -1  -1  -1  -1  -1 ZN  450 ASPA 327 kaaaaaakg-bbbbk
72 1ew2_A 316  116  3  0  2 DHH       Zn   4 112  -1  -1  -1  -1  -1 ZN 1001 ASPA 316 kaaaaaakg-bbbbk
73 1ah7_   55   67  4  1  0 DHHD      Zn  14  49   4  -1  -1  -1  -1 ZN  246 ASP   55 aakkb-akkb.-aaaaaaakk
74 1ak0_   45   75  4  0  0 DHHD      Zn  15  56   4  -1  -1  -1  -1 ZN  271 ASP   45 aaaaa-kkkbb-aaaaaaaak
75 1ush_   84  168  4  0  2 DNHH      Zn  32 101  35  -1  -1  -1  -1 ZN  601 ASP   84 kjgab-bbakk-bbda.-b..kg
76 1byf_A  52   74  2  0  2 DO        Zn  74  -1  -1  -1  -1  -1  -1 ZN  302 ASPA  52 gbbbb-.....
77 1lam_  255   79  4  1  1 DODE      Zn  77   0   2  -1  -1  -1  -1 ZN  488 ASP  255 abbkg-bkkdaaa
78 1ew2_A  42  315  3  0  1 DSD       Zn  50 265  -1  -1  -1  -1  -1 ZN 1002 ASPA  42 bbbgd-bbaaa-bba.b
79 1taf_B  31    4  2  1  0 ED        Zn   4  -1  -1  -1  -1  -1  -1 ZN 2001 GLUB  31 aaaaaaaaa
80 1noy_B 114  210  2  0  1 ED        Zn 210  -1  -1  -1  -1  -1  -1 ZN    2 GLUB 114 bbbbb-aaaaa
81 1d8w_B 234  100  4  0  0 EDHD      Zn  33  27  40  -1  -1  -1  -1 ZN  450 GLUB 234 bbbbb-bbbak-bbbb.-bbdbb
82 1qtw_A 145  116  4  1  0 EDHE      Zn  34  37  45  -1  -1  -1  -1 ZN  303 GLUA 145 bbbbb-bbbaa-.bbbb-bbdbk
83 3cao_A  26    3  2  0  0 EE        Zn   3  -1  -1  -1  -1  -1  -1 ZN  114 GLUA  26 aaaaaaak
84 1ec5_A  10   29  3  0  0 EEH       Zn  26   3  -1  -1  -1  -1  -1 ZN   50 GLUA  10 akaaaaaaaaaaaaaaagbaaaaaaaaaaaaaaa
85 1ete_C  58   22  2  1  0 EH        Zn  22  -1  -1  -1  -1  -1  -1 ZN 2135 GLUC  58 aaaaaaadbbaaaaaaaaaaaaakaaa
86 1cvr_A 152   59  2  1  0 EH        Zn  59  -1  -1  -1  -1  -1  -1 ZN  731 GLUA 152 bkbj.-b.b.b
87 1bf6_A 125   61  3  0  0 EHH       Zn  33  28  -1  -1  -1  -1  -1 ZN    2 GLUA 125 b.bbb-bbbbb-bbgkd
88 1dqs_B 194   93  3  0  2 EHH       Zn  77  16  -1  -1  -1  -1  -1 ZN  402 GLUB 194 aaaaa-bgkaa-bbaaa
89 1e4c_P  73   82  4  1  0 EHHH      Zn  19   2  61  -1  -1  -1  -1 ZN  999 GLUP  73 bkkaa-bbbbdba-bg..b
90 1a2p_C  60    2  2  1  0 EK        Zn   2  -1  -1  -1  -1  -1  -1 ZN  112 GLUC  60 gakgabb
91 1ctt_  102   30  3  1  0 HCC       Zn  27   3  -1  -1  -1  -1  -1 ZN  296 HIS  102 bbbaaaaaaaaaaakgba.ba.bbbkbbbbaaaaa
92 1btk_A 143   22  4  0  0 HCCC      Zn  11   1  10  -1  -1  -1  -1 ZN    1 HISA 143 bbbabbbbgg.baakgbkbakbbbkbb
93 1gpc_   64   26  4  0  0 HCCC      Zn  13  10   3  -1  -1  -1  -1 ZN    1 HIS   64 bbbbbbbggbbbbbbbkakdgdaakbaaaaa
94 1ptq_  231   49  4  0  0 HCCC      Zn  30   3  16  -1  -1  -1  -1 ZN    1 HIS  231 ...bb-bbbaakgb-ab...
95 1amp_   97   82  3  1  0 HDD       Zn  20  62  -1  -1  -1  -1  -1 ZN  502 HIS   97 .bbkb-aabda-bbb.k
96 1cg2_A 112   88  3  1  0 HDE       Zn  29  59  -1  -1  -1  -1  -1 ZN  501 HISA 112 b.bab-aabda-bdbbb
97 1j9y_A  79   42  3  1  0 HDE       Zn  32  10  -1  -1  -1  -1  -1 ZN 1003 HISA  79 bbgak-bbbkk-gakjb
```

```
 98 1i76_A 147    15   3   0   0 HDH        Zn    2  13  -1  -1  -1  -1  -1 ZN   998 HISA 147 .b.gkgdbbkjbjjba.bbb
 99 1vhh_  141    42   3   1   0 HDH        Zn    7  35  -1  -1  -1  -1  -1 ZN   400 HIS  141 bkakk-bbbbb-.bbbb
100 1lbu_  154    43   3   1   0 HDH        Zn    7  36  -1  -1  -1  -1  -1 ZN   214 HIS  154 bkakk-bbbbb-.bbbb
101 1qtw_A 182    49   3   2   0 HDH        Zn   47   2  -1  -1  -1  -1  -1 ZN   302 HISA 182 aaaaa-gbbbbbb
102 2cua_B 117     2   2   0   0 HE         Zn    2  -1  -1  -1  -1  -1  -1 ZN   172 HISB 117 bbbbbgb
103 1i1i_P 160     4   2   1   0 HE         Zn    4  -1  -1  -1  -1  -1  -1 ZN   702 HISP 160 baaaaaaaa
104 1ak0_   15   166   2   1   0 HE         Zn  166  -1  -1  -1  -1  -1  -1 ZN   274 HIS   15 aaakb-aaad.
105 2a0b_  728    30   3   0   0 HEE        Zn   26   4  -1  -1  -1  -1  -1 ZN   800 HIS  728 baaaaaaaaaakdkkbakaaakaaaaaaaaaakaa
106 1pyt_B  69   127   3   1   0 HEH        Zn    3 124  -1  -1  -1  -1  -1 ZN   350 HISB  69 gkbaadka-bb..b
107 2ctc_   69   127   3   1   0 HEH        Zn    3 124  -1  -1  -1  -1  -1 ZN   308 HIS   69 gabakdaa-bbbbb
108 1dq3_A 156     2   2   0   0 HH         Zn    2  -1  -1  -1  -1  -1  -1 ZN   901 HISA 156 bbbagbb
109 1fio_A  75     3   2   0   0 HH         Zn    3  -1  -1  -1  -1  -1  -1 ZN   502 HISA  75 aaaaaaaa
110 1hzy_A 201    29   2   2   1 HH         Zn   29  -1  -1  -1  -1  -1  -1 ZN   402 HISA 201 bbbbbaakgaaaaaaaaaakkgbbakkbbbbgak
111 1j79_A  16   234   3   1   1 HH  D      Zn    2  -1 232  -1  -1  -1  -1 ZN   400 HISA  16 bk.bbkj
112 1j98_A  54    72   3   1   0 HHC        Zn    4  68  -1  -1  -1  -1  -1 ZN   300 HISA  54 aaaaaaaaa-akbkd
113 1lba_   17   113   3   1   0 HHC        Zn  105   8  -1  -1  -1  -1  -1 ZN   151 HIS   17 bbbbk-bbaka-bkbkg
114 1ia9_A1751    63   4   0   0 HHCC       Zn   57   2   4  -1  -1  -1  -1 ZN  2001 HISA1751 aaaaa-aabbbbaaaaa
115 1hzy_B  55   246   3   1   1 HHD        Zn    2 244  -1  -1  -1  -1  -1 ZN   401 HISB  55 bkbbbab-bkgb.
116 1c7k_A  83    10   3   1   0 HHD        Zn    4   6  -1  -1  -1  -1  -1 ZN   133 HISA  83 aaaaaaaakgbbbda
117 1g12_A 117    13   3   2   0 HHD        Zn    4   9  -1  -1  -1  -1  -1 ZN   200 HISA 117 aaaaaaabaak.gbbbab
118 1ak0_  126    27   3   0   1 HHD        Zn   23   4  -1  -1  -1  -1  -1 ZN   272 HIS  126 akkdbabggggakbbbbbggbbbbaaaaaaaaa
119 1eu3_A 162    42   3   0   0 HHD        Zn   40   2  -1  -1  -1  -1  -1 ZN   401 HISA 162 bbbbk-ba.bbbb
120 1f0j_A 238   154   4   1   0 HHDD       Zn   36   1 117  -1  -1  -1  -1 ZN  1101 HISA 238 kaaaa-akkkbg-aaadk
121 1hs6_A 295    23   3   0   2 HHE        Zn    4  19  -1  -1  -1  -1  -1 ZN   701 HISA 295 aaaaaaakab.aabbbabkakaaaaaaa
122 1ezm_  140    24   3   1   0 HHE        Zn    4  20  -1  -1  -1  -1  -1 ZN   300 HIS  140 aaaaakaaaaak.agbbdbjaaaaaaaaa
123 1i1i_P 474    29   3   1   0 HHE        Zn    4  25  -1  -1  -1  -1  -1 ZN   701 HISP 474 aaaaaaaaaaaakbbbabakkb.kdbbakkakaa
124 1dmt_A 583    63   3   0   1 HHE        Zn    4  59  -1  -1  -1  -1  -1 ZN   755 HISA 583 aaaaaaakk-aaaaa
125 1hr6_B  70    80   3   1   0 HHE        Zn    4  76  -1  -1  -1  -1  -1 ZN   501 HISB  70 aaaaaaakk-aaaaa
126 1qtw_A  69    76   3   1   0 HHE        Zn   40  36  -1  -1  -1  -1  -1 ZN   301 HISA  69 dbbbb-bbba.-bbbbb
127 1bf6_A  12   231   4   0   0 HHED       Zn    2 111 118  -1  -1  -1  -1 ZN     1 HISA  12 bkbbbab-b.bbb-.agbk
128 1psz_A  67   213   4   0   0 HHED       Zn   72  66  75  -1  -1  -1  -1 ZN  1000 HISA  67 bkabb-bkdak-babdk-bakbb
129 1koq_A  92    19   3   1   0 HHH        Zn    2  17  -1  -1  -1  -1  -1 ZN   301 HISA  92 .bbbbbbabbbggbbbabbbbbb
130 2bc2_A  86    63   3   2   0 HHH        Zn    2  61  -1  -1  -1  -1  -1 ZN     1 HISA  86 ba.dbak-b.aba
```

```
131 1sml_A   84    76   3   1   0 HHH        Zn    2   74   -1   -1   -1   -1   -1 ZN   269 HISA   84 babdbak-b.aba
132 1i76_A 197    10   3   0   2 HHH        Zn    4    6   -1   -1   -1   -1   -1 ZN   999 HISA 197 aaaaaaaakgbbbbk
133 1lml_   264    70   3   1   0 HHH        Zn    4   66   -1   -1   -1   -1   -1 ZN   100 HIS   264 aaaaaaaak-gkbbb
134 1b66_B   23    27   3   0   1 HHH        Zn   25    2   -1   -1   -1   -1   -1 ZN   402 HISB   23 bb.bbbbkabbaaaaaaak.akkbakj.bbbb
135 1ton_    57    42   3   0   0 HHH        Zn   40    2   -1   -1   -1   -1   -1 ZN   200 HIS    57 baakb-.bbbkbg
136 1bkc_E 405    10   3   0   2 HHH        Zn    4    6   -1   -1   -1   -1   -1 ZN     1 HISE  405 aaaaaaaak.bbbbb
137 1qh5_B   54    80   4   1   0 HHHD      Zn    2   54   24   -1   -1   -1   -1 ZN   261 HISB   54 ba.kbak-b.abk-k.kab
138 1mfm_A   63    20   4   0   0 HHHD      Zn    8    9    3   -1   -1   -1   -1 ZN   154 HISA   63 jbbbdakgbbbbbakbabbbakbkj
139 1yai_A   70    21   4   0   0 HHHD      Zn    9    9    3   -1   -1   -1   -1 ZN   153 HISA   70 jbbbdakgagbb.bkkbadbbakbbb
140 1toa_A   68   211   4   0   0 HHHD      Zn   65   66   80   -1   -1   -1   -1 ZN   501 HISA   68 baakb-bkdkk-babdk-bakbb
141 1lam_   250    84   4   1   1 KDDE      Zn    5   18   61   -1   -1   -1   -1 ZN   489 LYS   250 bbbdbabbkg-kkkka-kdaaa
142 1b0n_A    1    63   3   1   0 #OE        Zn    0   63   -1   -1   -1   -1   -1 ZN 1004 META    1    .ka-kbkak
143 1lam_   173    98   2   1   0 OO         Zn   98   -1   -1   -1   -1   -1   -1 ZN   490 THR   173 aabba-kkkkk
144 1pmi_   111   174   4   1   0 QHEH      Zn    2   25  147   -1   -1   -1   -1 ZN   445 GLN   111 bbbbbbb-dbbbb-bb.bb
145 1taf_B   45     0   2   2   0 QQ         Zn    0   -1   -1   -1   -1   -1   -1 ZN 2007 GLNB   45 aaaaa
146 1j9y_A 208   112   4   0   0 RHDE      Zn    3   72   37   -1   -1   -1   -1 ZN 1004 ARGA 208 bbbbgbgk-b.bbb-babb.
147 1ak0_     1   119   4   0   0 #OHD      Zn    0    5  114   -1   -1   -1   -1 ZN   273 TRP     1    .baaaaaa-aaaak
148 1ca1_     1   129   4   1   0 #OHD      Zn    0   10  119   -1   -1   -1   -1 ZN   371 TRP     1    .b.-jaaaa-aaakk
149 1ah7_     1   121   4   1   0 #OHD      Zn    0   13  108   -1   -1   -1   -1 ZN   248 TRP     1    .bd-baaaa-aaakk
150 1fio_A   30     1   3   0   1 ##H        Zn    1    0   -1   -1   -1   -1   -1 ZN   501 META   30 ...aka
151 1qe3_A    2     1   3   2   1 ##H        Zn    1    0   -1   -1   -1   -1   -1 ZN   500 THRA    2  ..kgb
```

```
no   cngpname     nsp  np  nw  nn  dons      met  sd1  sd2  sd3  sd4  sd5  sd6  sd7  his   cn  pdb  cn2  rms  res   carbi    othdon   ecno
   1 1dps_K 140     2   2   2   0  OD          Na    2   -1   -1   -1   -1   -1   -1 ...    5 1dps   0 0.3 1.6   ...       |   ZZ                       DNA-BINDING PROT
   2 1e43_A 159    42   5   0   0  DDDDO      Na   24   11    6    1   -1   -1   -1 ..... 5 1e43   0 0.2 1.7   ..b..     |          3.2.1.1;          HYD       ROLASE
   3 1hx6_A 143     6   3   3   0  DND        Na    3    3   -1   -1   -1   -1   -1 ...    6 1hx6   0 0.1 1.6   ...       |   ZZZ                       VIRUS/VIRAL PROT
   4 1dp0_A 201   403   3   2   0  DON        Na  400    3   -1   -1   -1   -1   -1 ...    5 1dp0   0 0.2 1.7   ...       |   ZZ    3.2.1.23;         HYD         ROLASE
   5 1g29_1  82    75   3   0   0  DOO        Na   74    1   -1   -1   -1   -1   -1 ...    3 1g29   0 0.6 1.9   ...       |                           SUGAR BINDING PR
   6 1qus_A 237    14   5   0   0  DSDOD      Na    2    2    2    8   -1   -1   -1 ... . 5 1qus   1 0.1 1.7   ....b     |          3.2.1.-;          HYD   ROLASE
   7 1hx6_B 111   150   2   4   0  EE         Na  150   -1   -1   -1   -1   -1   -1 ..     6 1hx6   0 0.2 1.6   ..        | ZZZZ                        VIRUS/VIRAL PROT
   8 1g29_1  11    42   2   1   0  KE         Na   42   -1   -1   -1   -1   -1   -1 ..     3 1g29   0 0.2 1.9   ..        | Z                           SUGAR BINDING PR
   9 1hzy_A  38   116   2   4   0  NO         Na  116   -1   -1   -1   -1   -1   -1 ..     6 1hzy   1 0.5 1.3   ..        | ZZZZ   3.1.8.1;          HYD       ROLASE
  10 1eyz_A 100     3   3   3   0  NOO        Na    1    2   -1   -1   -1   -1   -1 ...    6 1eyz   0 0.3 1.8   ...       |   ZZZ   2.1.2.-;          TRA   NSFERASE
  11 1dps_A 140     2   2   3   0  OD         Na    2   -1   -1   -1   -1   -1   -1 ..     5 1dps   0 0.3 1.6   ..        | ZZZ                          DNA-BINDING PROT
  12 1svy_  187    23   3   1   0  ODE        Na    1   22   -1   -1   -1   -1   -1 ...    4 1svy   1 0.2 1.8   ..b       |   Z                          ACTIN-BINDING PR
  13 1gof_   29   113   7   0   0  ODOOTOE   Na    3    2    3    0  104    1   -1 ......7 1gof   0 0.3 1.7   ....... |         1.1.3.9) (       OXI       DOREDUCTASE(O
  14 1g29_2 181     0   2   0   0  OE         Na    0   -1   -1   -1   -1   -1   -1 ..     2 1g29   0 0.2 1.9   ..        |                          SUGAR BINDING PR
  15 1et1_B   1     3   2   0   0  OE         Na    3   -1   -1   -1   -1   -1   -1 ..     2 1et1   0 0.5 0.9   .b        |                      HORMONE/GROWTH F
  16 1cm5_B 654    47   3   0   0  OEO        Na   46    1   -1   -1   -1   -1   -1 ...    3 1cm5   1 0.2 2.3   ...       |          2.3.1.54;        TRA    NSFERASE
  17 2occ_A  40   401   4   0   0  OEOO       Na    0    5  396   -1   -1   -1   -1 ....   4 2occ   0 0.1 2.3   ....      |         1.9.3.1;          OXI   DOREDUCTASE
  18 1ebf_B 143     7   5   0   0  OEOOO      Na    0    3    2    2   -1   -1   -1 ..... 5 1ebf   0 0.3 2.3   .....     |         1.1.1.3;          OXI DOREDUCTASE
  19 1hx6_A 262     0   2   3   0  ON         Na    0   -1   -1   -1   -1   -1   -1 ..     5 1hx6   0 0.2 1.6   ..        | ZZZ                         VIRUS/VIRAL PROT
  20 1c9o_A  20     3   2   4   0  OO         Na    3   -1   -1   -1   -1   -1   -1 ..     6 1c9o   1 0.1 1.2   ..        | ZZZZ                         TRANSCRIPTION
  21 1cm5_A 652    49   4   0   0  OOEO       Na    2   46    1   -1   -1   -1   -1 ....   4 1cm5   0 0.1 2.3   ....      |          2.3.1.54;        TRA       NSFERASE
  22 1bpy_A  60     5   3   2   1  OOO        Na    2    3   -1   -1   -1   -1   -1 ...    6 1bpy   0 0.4 2.2   ...       |   UZZ   2.7.7.7;          COM     PLEX (NUCLEOT
  23 1bpy_A 101     5   3   2   1  OOO        Na    2    3   -1   -1   -1   -1   -1 ...    6 1bpy   0 0.5 2.2   ...       |   UZZ   2.7.7.7;          COM      PLEX (NUCLEOT
  24 1c24_A  74   157   3   1   0  OOO        Na    2  155   -1   -1   -1   -1   -1 ...    4 1c24   1 0.4 1.7   ...       |   Z     3.4.11.18;       HYD        ROLASE
  25 1dp0_A 556     6   3   2   0  OOO        Na    3    3   -1   -1   -1   -1   -1 ...    5 1dp0   1 0.2 1.7   ...       |   ZZ    3.2.1.23;         HYD     ROLASE
  26 1g3k_A 157     6   3   3   0  OOO        Na    3    3   -1   -1   -1   -1   -1 ...    6 1g3k   0 0.2 1.9   ...       |   ZZZ   3.4.99.-;         HYD      ROLASE
  27 1dp0_B 932    38   3   1   0  OOO        Na   35    3   -1   -1   -1   -1   -1 ...    4 1dp0   0 0.3 1.7   ...       |   Z     3.2.1.23;         HYD      ROLASE
  28 1f6d_C 298    54   3   0   0  OOO        Na   52    2   -1   -1   -1   -1   -1 ...    3 1f6d   2 0.5 2.5   ...       |          5.1.3.14;        ISO MERASE
  29 1qop_B 232    76   3   2   0  OOO        Na   74    2   -1   -1   -1   -1   -1 ...    5 1qop   0 0.1 1.4   ...       |   ZZ    4.2.1.20;         LYA       SE
  30 1dp0_B 647    23   3   2   1  OOO        Na    3   20   -1   -1   -1   -1   -1 ... . 7 1dp0   0 0.4 1.7   ... .     |   X  ZZ 3.2.1.23;        HYD          ROLASE
  31 1lla_  507    71   4   2   0  OOOD       Na    3   68    0   -1   -1   -1   -1 ....   6 1lla   0 0.3 2.2   ....      |    ZZ                        OXYGEN TRANSPORT
```

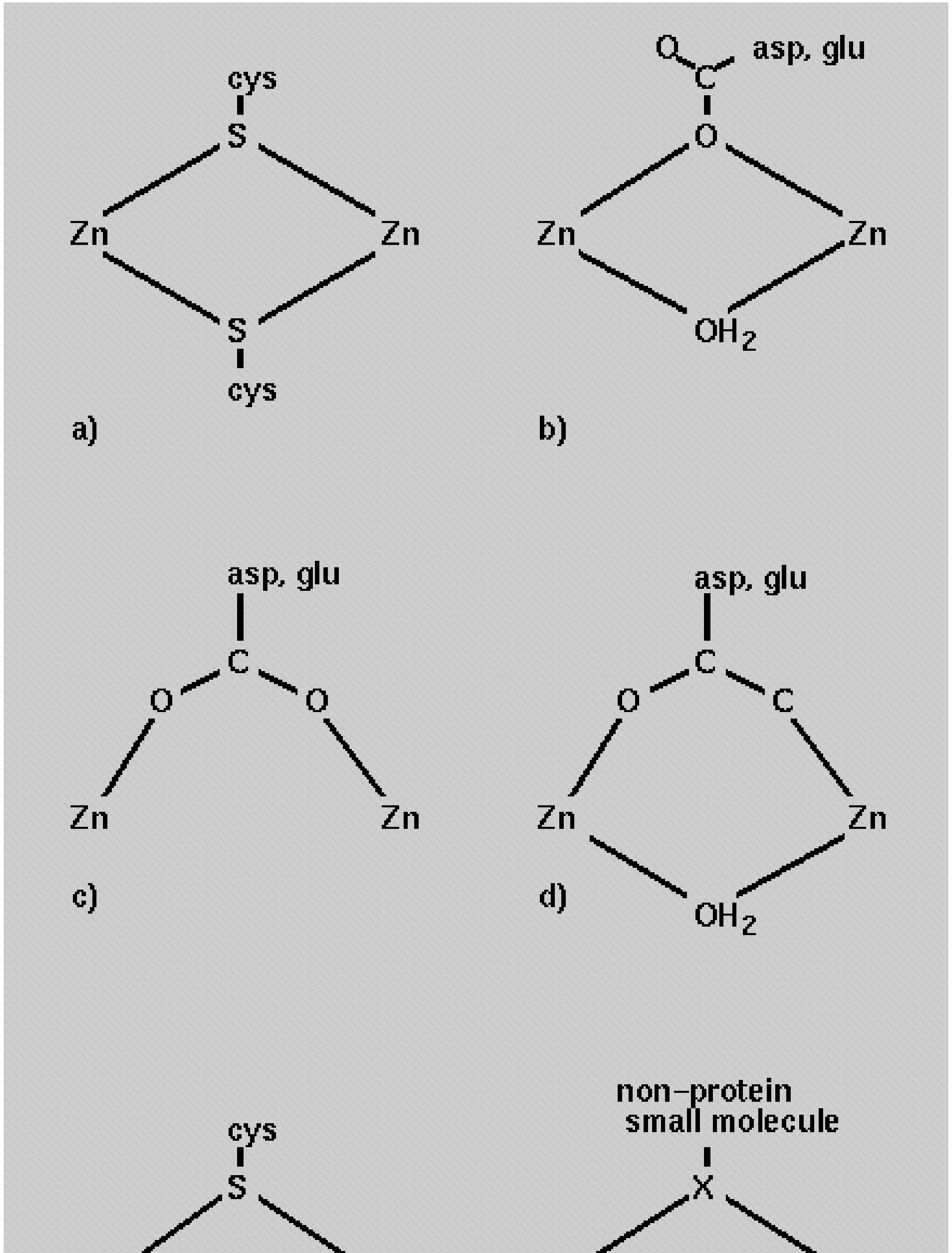
```
32 1dp0_C 556     6   4   2   0 OOOO      Na    3    1    2   -1   -1   -1   -1 ....   6 1dp0   0 0.4 1.7   ....     |    ZZ   3.2.1.23;        HYD       ROLASE
33 1gen_   478   142   4   1   1 OOOO      Na   45   48   49   -1   -1   -1   -1 ....   6 1gen   0 0.3 2.2   ....     |    JZ   3.4.24.24;       HYD           ROLASE (METAL
34 1hxn_   236   140   4   1   0 OOOO      Na   45   52   43   -1   -1   -1   -1 ....   5 1hxn   0 0.3 1.8   ....     |    Z                      BINDING PROTEIN
35 1e39_A 506    30   5   1   0 OOOOO     Na    1    1   26    2   -1   -1   -1 ..... 6 1e39   0 0.3 1.8   .....    |     Z                     FUM  ARATE REDUCTA
36 1gdo_A 154    20   3   2   0 OOS       Na   20    0   -1   -1   -1   -1   -1 ...    5 1gdo   0 0.2 1.8   ...      |   ZZ    2.6.1.16;        COM       PLEX (TRANSFE
37 1d7u_A   95     7   5   0   0 OOTOO     Na    3    0    1    3   -1   -1   -1 ..... 5 1d7u   0 0.3 2.0   .....    |          4.1.1.64;        LYA    SE
38 1pox_B 452     3   2   1   0 OQ         Na    3   -1   -1   -1   -1   -1   -1 ..     3 1pox  20  0.4 2.1   ..       | Z      1.2.3.3) M      OXI  DOREDUCTASE(O
39 1ewn_A 149    28   5   1   0 OSOOO     Na   22    1    2    3   -1   -1   -1 ..... 6 1ewn   1 0.1 2.1   .....    |     Z   3.2.2.21;       HYD           ROLASE/DNA
40 1cfb_   679    64   4   1   0 SOON      Na    1   60    3   -1   -1   -1   -1 ....   5 1cfb   0 0.3 2.0   ....     |    Z                      RANEUADHESION
```

```
no   cngpname     nsp  np  nw  nn  dons     met  sd1  sd2   sd3  sd4  sd5  sd6  sd7  metal    startaa    |  conf_efimov

   1  1dps_K  140      2   2   2   0  OD         Na    2   -1   -1   -1   -1   -1   -1  NA     11  LYSK  140  kbabaaa

   2  1e43_A  159     42   5   0   0  DDDDO      Na   24   11    6    1   -1   -1   -1  NA    505  ASPA  159  kba.b-.bbbb-bkabab.bbbbk

   3  1hx6_A  143      6   3   3   0  DND        Na    3    3   -1   -1   -1   -1   -1  NA    705  ASPA  143  b.bakgabb.b

   4  1dp0_A  201    403   3   2   0  DON        Na  400    3   -1   -1   -1   -1   -1  NA   3101  ASPA  201  dbb.b-.kkkbkba

   5  1g29_1   82     75   3   0   0  DOO        Na   74    1   -1   -1   -1   -1   -1  NA    422  ASP1   82  kbdbb-aakgba

   6  1qus_A  237     14   5   0   0  DSDOD      Na    2    2    2    8   -1   -1   -1  NA    400  ASPA  237  bbbkagkgbbbkkbaaaaa

   7  1hx6_B  111    150   2   4   0  EE         Na  150   -1   -1   -1   -1   -1   -1  NA    706  GLUB  111  .bbbb-bbbbb

   8  1g29_1   11     42   2   1   0  KE         Na   42   -1   -1   -1   -1   -1   -1  NA    112  LYS1   11  bbbbb-gabbb

   9  1hzy_A   38    116   2   4   0  NO         Na  116   -1   -1   -1   -1   -1   -1  NA    405  ASNA   38  bbbbk-aaaaa

  10  1eyz_A  100      3   3   3   0  NOO        Na    1    2   -1   -1   -1   -1   -1  NA    961  ASNA  100  gbbbba.a

  11  1dps_A  140      2   2   3   0  OD         Na    2   -1   -1   -1   -1   -1   -1  NA      1  LYSA  140  kbabaaa

  12  1svy_   187     23   3   1   0  ODE        Na    1   22   -1   -1   -1   -1   -1  NA      2  GLY   187  bkkbbbbbbk.kbbbbbbbakbbaaaaa

  13  1gof_    29    113   7   0   0  ODOOTOE    Na    3    2    3    0  104    1   -1  NA    702  LYS    29  akakdagbkabbb-bbabbb

  14  1g29_2  181      0   2   0   0  OE         Na    0   -1   -1   -1   -1   -1   -1  NA    265  GLU2  181  aaaaa

  15  1et1_B    1      3   2   0   0  OE         Na    3   -1   -1   -1   -1   -1   -1  NA    101  SERB    1    .kaaaa

  16  1cm5_B  654     47   3   0   0  OEO        Na   46    1   -1   -1   -1   -1   -1  NA   1057  LEUB  654  aabba-kbajba

  17  2occ_A   40    401   4   0   0  OEOO       Na    0    5  396   -1   -1   -1   -1  NA    519  GLUA   40  aaakkbbjba-gbkbb

  18  1ebf_B  143      7   5   0   0  OEOOO      Na    0    3    2    2   -1   -1   -1  NA   2114  GLUB  143  bdakabg.kbda

  19  1hx6_A  262      0   2   3   0  ON         Na    0   -1   -1   -1   -1   -1   -1  NA    703  ASNA  262  bbbbb

  20  1c9o_A   20      3   2   4   0  OO         Na    3   -1   -1   -1   -1   -1   -1  NA    101  VALA   20  bbbbgbbb

  21  1cm5_A  652     49   4   0   0  OOEO       Na    2   46    1   -1   -1   -1   -1  NA   1056  ALAA  652  aaaabba-kbajba

  22  1bpy_A   60      5   3   2   1  OOO        Na    2    3   -1   -1   -1   -1   -1  NA    342  LYSA   60  aaakbbgbba

  23  1bpy_A  101      5   3   2   1  OOO        Na    2    3   -1   -1   -1   -1   -1  NA    341  THRA  101  aakkbbgbba

  24  1c24_A   74    157   3   1   0  OOO        Na    2  155   -1   -1   -1   -1   -1  NA    501  ASNA   74  bb.kba.-bbbbb

  25  1dp0_A  556      6   3   2   0  OOO        Na    3    3   -1   -1   -1   -1   -1  NA   3102  PHEA  556  aaaaabakbkj

  26  1g3k_A  157      6   3   3   0  OOO        Na    3    3   -1   -1   -1   -1   -1  NA    500  GLYA  157  aaakabkkbbk

  27  1dp0_B  932     38   3   1   0  OOO        Na   35    3   -1   -1   -1   -1   -1  NA   3103  PROB  932  abb.b-aaakkbkb

  28  1f6d_C  298     54   3   0   0  OOO        Na   52    2   -1   -1   -1   -1   -1  NA   3378  PROC  298  aakkk-...kbbb

  29  1qop_B  232     76   3   2   0  OOO        Na   74    2   -1   -1   -1   -1   -1  NA    501  GLYB  232  bbaag-dkbk..b

  30  1dp0_B  647     23   3   2   1  OOO        Na    3   20   -1   -1   -1   -1   -1  NA   3104  SERB  647  bbdkdbbbbbbbbbggbba...bbbbdb

  31  1lla_   507     71   4   2   0  OOOD       Na    3   68    0   -1   -1   -1   -1  NA    631  SER   507  akbkkbba-gkada
```

```
 32  1dp0_C  556      6   4   2   0  OOOO       Na    3    1    2   -1   -1   -1   -1  NA  3102  PHEC  556  aaaaabakbkj
 33  1gen_   478   142   4   1   1  OOOO       Na   45   48   49   -1   -1   -1   -1  NA   304  ILE   478  abbbb-a.bbb-abbbb-a.bbb
 34  1hxn_   236   140   4   1   0  OOOO       Na   45   52   43   -1   -1   -1   -1  NA     4  SER   236  bbk.b-bba.b-bbabb-.ba.b
 35  1e39_A  506    30   5   1   0  OOOOO      Na    1    1   26    2   -1   -1   -1  NA   810  THRA  506  .bbab.
bbbbkkgbbbbkkgbbbbgbbbbjkabbg
 36  1gdo_A  154    20   3   2   0  OOS        Na   20    0   -1   -1   -1   -1   -1  NA   300  GLYA  154  bb.bbbbbbbbkkdkkbbbbbbdgb
 37  1d7u_A   95      7   5   0   0  OOTOO      Na    3    0    1    3   -1   -1   -1  NA   435  ALAA   95  aaaakbbbgbab
 38  1pox_B  452      3   2   1   0  OQ          Na    3   -1   -1   -1   -1   -1   -1  NA   614  METB  452  aaakaaka
 39  1ewn_A  149    28   5   1   0  OSOOO      Na   22    1    2    3   -1   -1   -1  NA   501  META  149  aaakbbbgbbbbbbbgg.bbbbbbbbb.bgbbb
 40  1cfb_   679    64   4   1   0  SOON       Na    1   60    3   -1   -1   -1   -1  NA     5  SER   679  bbbbg.-bakkag.b
```

```
no   cngpname     nsp  np  nw  nn  dons     met  sd1  sd2  sd3  sd4  sd5  sd6  sd7  his   cn  pdb  cn2  rms  res   carbi    othdon   ecno

   1 1bup_A   10      5   2   2   3  DO          K      5   -1   -1   -1   -1   -1   -1 ..      7  1bup   1  0.3  1.7   b.        |  UUUZZ                       HYDROLASE

   2 1a9x_E4841      2   2   0   2  EN          K      2   -1   -1   -1   -1   -1   -1 ..      4  1a9x   0  0.3  1.8   b.        |  UU                          AMIDOTRANSFERASE

   3 1a9x_A  215    32   6   0   0  ENOOOS    K    21    2    1    3    5   -1   -1 ......6  1a9x   0  0.2  1.8   ......   |                               AMIDOTRANSFERASE

   4 1a9x_A  761    31   5   0   0  EOOOS     K    22    1    3    5   -1   -1   -1 .....  5  1a9x   0  0.2  1.8   .....    |                               AMIDOTRANSFERASE

   5 1a9x_C2217    68   5   2   1  ETNNQ     K    27   39    0    2   -1   -1   -1 .....  8  1a9x   0  0.5  1.8   .....    |     UZZ                     AMI    DOTRANSFERASE

   6 1a9x_E4217    68   4   1   1  ETNQ      K    27   39    2   -1   -1   -1   -1 ....   6  1a9x   0  0.4  1.8   ....     |    UZ                        AMIDOTRANSFERASE

   7 1jf8_A   13    52   5   2   0  NOSOD     K    23    0   27    2   -1   -1   -1 .....  7  1jf8   0  0.2  1.1   .....    |     ZZ                       OXIDOREDUCTASE

   8 1bup_A  199     7   5   1   1  ODOTD     K     0    5    0    2   -1   -1   -1 .....  7  1bup   0  0.3  1.7   .....    |     UZ                       HYDROLASE

   9 1a9x_G6126   174   4   3   0  OEEO      K     1  172    1   -1   -1   -1   -1 ....   7  1a9x   1  0.3  1.8   ....     |    ZZZ                      AMI   DOTRANSFERASE

  10 1a9x_C2126   175   5   3   0  OEEON     K     1  172    1    1   -1   -1   -1 .....  8  1a9x   0  0.3  1.8   .....    |     ZZZ                    AMI      DOTRANSFERASE

  11 1k4c_C   77      1   2   0   1  OO          K      1   -1   -1   -1   -1   -1   -1 ..      3  1k4c  40  0.3  2.0   ..        |  J                          MEMBRANE PROTEIN

  12 1k4c_C   76      1   2   0   2  OO          K      1   -1   -1   -1   -1   -1   -1 ..      4  1k4c  40  0.3  2.0   ..        |  JJ                         MEMBRANE PROTEIN

  13 1k4c_C   75      1   2   0   2  OO          K      1   -1   -1   -1   -1   -1   -1 ..      4  1k4c  40  0.3  2.0   ..        |  JJ                         MEMBRANE PROTEIN

  14 1guq_A  153    13   2   5   0  OO          K    13   -1   -1   -1   -1   -1   -1 ..      7  1guq   0  0.2  1.8   ..        |  ZZZZZ  2.7.7.10;         NUC         LEOTIDYLTRANS

  15 1a9x_D3516    96   2   2   0  OO          K    96   -1   -1   -1   -1   -1   -1   .      4  1a9x   0  0.2  1.8   ..        |  ZZ                          AMIDOTRANSFERASE

  16 1ad1_B   75      4   3   2   0  OOO        K      2    2   -1   -1   -1   -1   -1 ...     5  1ad1   0  0.4  2.2   ...       |   ZZ    2.5.1.15;         TRA  NSFERASE

  17 1aop_   362    35   4   3   0  OOON       K    33    1    1   -1   -1   -1   -1 ....   7  1aop   0  0.3  1.6   ....     |    ZZZ  1.8.1.2;          OXI     DOREDUCTASE

  18 2irf_J  685     6   4   0   0  OOOO       K     1    2    3   -1   -1   -1   -1 ....   4  2irf  20  0.1  2.2   ....     |                            E  REGULATION/ GEN

  19 1elu_A  316     6   4   1   0  OOOO       K     1    2    3   -1   -1   -1   -1 ....   5  1elu   0  0.3  1.5   ....     |    Z                        SE LYA

  20 1g8m_A  426   164   6   0   0  OOSSDO    K     3    2    2  107   50   -1   -1 ......6  1g8m   0  0.2  1.8   ......   |         2.1.2.3, 3       TRA              NSFERASE, HYD

  21 1a9x_E4084    30   3   2   0  OOT        K    28    2   -1   -1   -1   -1   -1 ...     5  1a9x   0  0.3  1.8   ...       |   ZZ                         AMIDOTRANSFERASE

  22 1a9x_C2084    30   3   1   0  OOT        K    28    2   -1   -1   -1   -1   -1 ....   5  1a9x   0  0.5  1.8   ....     |    Z                         AMIDOTRANSFERASE

  23 1eex_A  261    19   4   2   0  OSEE       K     3    1   15   -1   -1   -1   -1 ....   6  1eex   0  0.2  1.7   ..b.     |    ZZ   4.2.1.28;         LYA       SE

  24 1d7u_A   78   229   5   1   0  OSOOD     K     2  223    2    2   -1   -1   -1 .....  6  1d7u   0  0.2  2.0   .....    |     Z   4.1.1.64;        LYA         SE

  25 1k4c_C   75      0   2   0   1  OT          K      0   -1   -1   -1   -1   -1   -1 ..      3  1k4c  40  0.3  2.0   ..        |  J                          MEMBRANE PROTEIN

  26 1a9x_E4143     1   3   3   0  OTO        K     0    1   -1   -1   -1   -1   -1 ...     6  1a9x   0  0.3  1.8   ...       |   ZZZ                        AMIDOTRANSFERASE

  27 1jdr_A  176    23   7   0   0  OTTONOD   K     0   16    2    0    3    2   -1 ......7  1jdr   0  0.2  1.5   ....... |          1.11.1.5;        OXI      DOREDUCTASE

  28 1eex_A  141   221   5   0   2  QEEQO     K    29   51   75   66   -1   -1   -1 ..... 7  1eex   0  0.5  1.7   .....    |     UU  4.2.1.28;        LYA              SE
```

```
no   cngpname     nsp  np  nw  nn  dons     met  sd1  sd2   sd3  sd4  sd5  sd6  sd7  metal    startaa    |  conf_efimov
   1  1bup_A   10      5   2   2   3  DO          K      5   -1   -1   -1   -1   -1   -1   K   490  ASPA   10  bbbbbaabbb
   2  1a9x_E4841      2   2   0   2  EN          K      2   -1   -1   -1   -1   -1   -1   K  5912  GLUE4841  babbbd.
   3  1a9x_A  215    32   6   0   0  ENOOOS    K    21    2    1    3    5   -1   -1   K  1903  GLUA  215  bbbbb-bbbabbbgbbaakbbb
   4  1a9x_A  761    31   5   0   0  EOOOS     K    22    1    3    5   -1   -1   -1   K  1913  GLUA  761  bbbbb-babbbgbbakkbbb
   5  1a9x_C2217    68   5   2   1  ETNNQ     K    27   39    0    2   -1   -1   -1   K  3942  GLUC2217  bbbbb-bbakk-bbbbbbb
   6  1a9x_E4217    68   4   1   1  ETNQ      K    27   39    2   -1   -1   -1   -1   K  5942  GLUE4217  bbbbb-bbakk-bbbbbbb
   7  1jf8_A   13    52   5   2   0  NOSOD     K    23    0   27    2   -1   -1   -1   K   133  ASNA   13  akga.-bbbbb-kbbbbbb
   8  1bup_A  199     7   5   1   1  ODOTD     K     0    5    0    2   -1   -1   -1   K   491  ASPA  199  bbbb.akbbbbb
   9  1a9x_G6126   174   4   3   0  OEEO      K     1  172    1   -1   -1   -1   -1   K  7904  ALAG6126  aaaaba-ba.bbd
  10  1a9x_C2126   175   5   3   0  OEEON     K     1  172    1    1   -1   -1   -1   K  3904  ALAC2126  aaaaba-ba.bbd.
  11  1k4c_C   77      1   2   0   1  OO          K      1   -1   -1   -1   -1   -1   -1   K  3001  GLYC   77  gagaga
  12  1k4c_C   76      1   2   0   2  OO          K      1   -1   -1   -1   -1   -1   -1   K  3002  VALC   76  kgagag
  13  1k4c_C   75      1   2   0   2  OO          K      1   -1   -1   -1   -1   -1   -1   K  3003  THRC   75  akgaga
  14  1guq_A  153    13   2   5   0  OO         K    13   -1   -1   -1   -1   -1   -1   K   353  ASNA  153  bbbb.akkgbkbkbb.bb
  15  1a9x_D3516    96   2   2   0  OO         K    96   -1   -1   -1   -1   -1   -1   K  3941  HISD3516  bbb.b-bbgbb
  16  1ad1_B   75      4   3   2   0  OOO        K      2    2   -1   -1   -1   -1   -1   K      1  VALB   75  akakbd.bb
  17  1aop_   362    35   4   3   0  OOON       K    33    1    1   -1   -1   -1   -1   K   590  ILE   362  bbbbg-bkkgbbb
  18  2irf_J  685     6   4   0   0  OOOO       K     1    2    3   -1   -1   -1   -1   K  4005  METJ  685  aaaaabakbbb
  19  1elu_A  316     6   4   1   0  OOOO       K     1    2    3   -1   -1   -1   -1   K  1001  LEUA  316  aaaaabakbbb
  20  1g8m_A  426   164   6   0   0  OOSSDO    K     3    2    2  107   50   -1   -1   K  1001  VALA  426  aaaakbbbbbbb-bbkbb-bbdbb
  21  1a9x_E4084    30   3   2   0  OOT        K    28    2   -1   -1   -1   -1   -1   K  5940  ASPE4084  gbabbbkkkd.aaaaaaaaaaaakgaaaakgbbbb
  22  1a9x_C2084    30   3   1   0  OOT        K    28    2   -1   -1   -1   -1   -1   K  3940  ASPC2084  gbabbbkkkdjaaaaaaaaaaaakgaaaakgbbbb
  23  1eex_A  261    19   4   2   0  OSEE       K     3    1   15   -1   -1   -1   -1   K   604  GLYA  261  b..bgaaaakgbkggbbaaaaaaa
  24  1d7u_A   78   229   5   1   0  OSOOD     K     2  223    2    2   -1   -1   -1   K   436  LEUA   78  abbbkkb-bbakbgbaa
  25  1k4c_C   75      0   2   0   1  OT          K      0   -1   -1   -1   -1   -1   -1   K  3004  THRC   75  akgag
  26  1a9x_E4143     1   3   3   0  OTO        K     0    1   -1   -1   -1   -1   -1   K  5943  THRE4143  bbbbbb
  27  1jdr_A  176    23   7   0   0  OTTONOD   K     0   16    2    0    3    2   -1   K   500  THRA  176  kkkbj.bbaaakgbbjbabadkkbbbka
  28  1eex_A  141   221   5   0   2  QEEQO     K    29   51   75   66   -1   -1   -1   K   603  GLNA  141  bbbbb-bbbb.-.badb-bbbbg-abd.b
```


cys
S
Zn
Zn
S
cys
a)
O
C
asp, glu
O
Zn
Zn
OH2
b)
asp, glu
C
O
O
Zn
Zn
c)
asp, glu
C
O
C
Zn
Zn
OH2
d)
cys
S
non-protein
small molecule
X

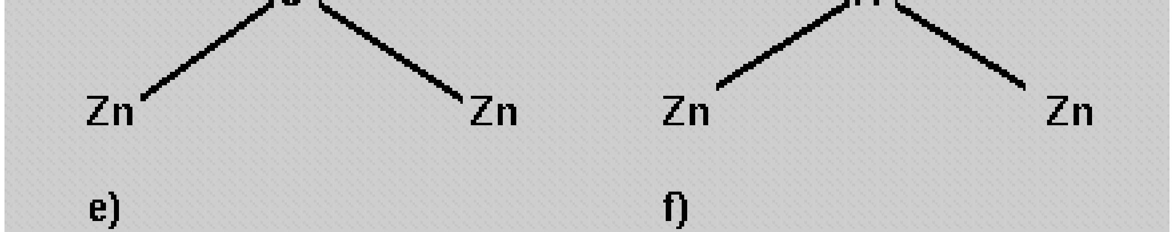


Fig. 5

```
   statistics for             190 Ca coordination groups
   ==================================================


                                    N:    0    1    2    3    4    5    6    7    8
   Numbers with
    total coordn number = N        0    0    2    6   13   36  110   22    1
    N protein donor groups         0    0   29   26   45   61   27    2    0
    N non-protein donors          51   60   39   26    8    5    1    0    0


   (Note: total coord.no includes both Os of bidentate carboxlyate;
   numbers of protein and non protein donors treat a -COO group as one donor.
         also, reminder:  coordination nos >8 not recorded correctly by makseq)


   Numbers of cngps with span, min and max seqdif =
       0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000
 2  protein donors in cngp
      4   3   5   4   4   0   0   0   0   0   1    0    3    1    0    3    0    1    0    0    0    0 span
      4   3   5   4   4   0   0   0   0   0   1    0    3    1    0    3    0    1    0    0    0    0 min
 3  protein donors in cngp
      0   1   0   0   2   3   0   1   1   1   0    2    5    2    3    1    2    1    0    1    0    0 span
      2   4  12   3   1   1   0   0   0   1   0    1    0    1    0    0    0    0    0    0    0    0 min
      0   1   1   4   0   2   1   0   0   0   1    2    6    2    1    1    2    1    1    0    0    0 max
 4  protein donors in cngp
      0   0   0   0   0   1   2   1   2   2   3    1    4    2    5   13    4    1    3    1    0    0 span
      3   6  26   2   2   1   0   0   0   2   0    2    0    0    1    0    0    0    0    0    0    0 min
      0   0   2   3   1   5   0   0   0   1   0    1    6    5    9    6    3    0    2    1    0    0 max
 5  protein donors in cngp
      0   0   0   0   0   0   0   0   1   0   4   27    7    3    3    9    6    1    0    0    0    0 span
      7  15  37   2   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
      0   0   1   4   0  25   0   0   1   0   0    6    5    4    2    7    5    1    0    0    0    0 max
 6  protein donors in cngp
      0   0   0   0   0   0   0   0   1   1   0    5    7    2    1    8    0    1    1    0    0    0 span
```

```
       8   7 12   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
       0   0   0   4   0   1   1   0   0   0   0    7    2    8    0    2    0    2    0    0    0    0 max
  7  protein donors in cngp
       0   0   0   0   0   0   0   0   0   0   0    0    1    0    0    1    0    0    0    0    0    0 span
       2   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
       0   0   0   0   0   0   0   0   0   0   0    1    0    0    0    1    0    0    0    0    0    0 max


     numbers of each type of donor
       D     E     S     T     O     H     C     K     M     N     Q     R     Y     W     #
     269   109    18    15   305     3     0     0     0    64    13     0     1     0     0


     numbers of donor pairs found with each seqdif
     ---------------------------------------------
        0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000


    DD   0   3  35   4   3   1   2   0   0   1   0    3    2    1    1    4    2    0    0    0    0    0
    DE   0   3   1   2   2   0   0   1   0   0   0    0    1    2    1    2    1    0    0    0    0    0
    DS   0   0   6   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
    DT   0   0   3   0   1   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
    DO   0  19  38   7   1   2   0   0   0   0   0    7    3   13    0    2    2    2    1    0    0    0
    DN   0   4  16   0   0   1   0   0   0   0   0    0    2    1    0    0    0    0    0    0    0    0
    DQ   0   0   0   0   0   0   0   0   0   0   1    0    0    0    0    0    0    0    0    0    0    0
    ED   0   0   4   0   1   0   0   0   1   0   0    2    4    1    0    1    0    0    1    0    0    0
    EE   0   1   2   6   1   0   0   0   0   0   0    0    0    0    0    2    0    0    0    0    0    0
    EO   0   3   4   1   0   0   0   0   0   0   1    1    3    2    1    1    0    0    0    0    0    0
    EH   0   0   0   0   0   1   1   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
    EN   0   0   1   1   0   0   0   0   0   0   0    2    0    0    0    1    0    0    0    0    0    0
    SD   0   0   2   0   0   0   0   0   0   0   0    1    0    0    1    0    0    0    0    0    0    0
    SE   0   0   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
    SS   0   0   0   0   0   0   0   0   0   0   0    0    0    1    0    0    0    0    0    0    0    0
    SO   0   2   3   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
    SN   0   0   0   0   0   0   0   0   0   0   0    0    0    1    0    0    0    0    0    0    0    0
    SQ   0   0   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
```

```
  TD   0   0   1   0   0   0   0   0   0   0   0    0    1    0    0    1    0    0    0    0    0    0

  TO   0   0   2   2   0   0   0   0   0   0   0    0    0    0    0    0    0    1    0    0    0    0

  TN   0   0   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  TQ   0   0   0   1   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  OD  12   5  32  12   2   2   0   2   1   2   1    4    2    1    1    0    2    0    0    1    0    0

  OE   1   1   8   1   0  27   0   0   0   2   0    0    0    0    1    0    0    1    0    0    0    0

  OS   5   0   3   1   0   0   0   0   0   0   0    0    1    0    0    0    0    0    0    0    0    0

  OT   7   0   3   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  OO   0   5  38  20   2   2   0   0   0   0   0    2    5    4    5    2    3    1    0    0    0    0

  OH   1   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  ON   2   1   6   6   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  OQ   1   0   4   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  OY   1   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  HD   0   0   0   1   0   0   0   0   0   0   0    0    0    0    0    0    1    0    0    0    0    0

  ND   0   4   7   0   0   1   0   0   0   0   0    1    1    2    1    0    0    0    0    0    0    0

  NE   0   0   0   2   0   0   1   0   1   0   0    0    2    0    0    2    0    0    0    0    0    0

  NO   0   3  15   0   0   0   0   0   0   0   0    2    1    1    2    3    0    0    0    0    0    0

  NN   1   2   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  QD   0   0   0   0   0   0   0   0   0   0   0    1    0    4    0    0    0    0    0    0    0    0

  QE   0   0   0   1   1   1   0   1   0   0   0    0    0    0    0    1    0    0    0    0    0    0


 all  31  56237  68  14  38   4   4   3   5   3   26   28   34   14   22   11    5    2    1    0    0
```

```
   statistics for             81 Mg coordination groups
   ==================================================


                                N:    0    1    2    3    4    5    6    7    8
   Numbers with
    total coordn number = N        0    0    3    8    7   16   46    1    0
    N protein donor groups         0    0   41   31    6    3    0    0    0
    N non-protein donors          10   12    9   25   24    1    0    0    0

   (Note: total coord.no includes both Os of bidentate carboxlyate;
   numbers of protein and non protein donors treat a -COO group as one donor.
         also, reminder:  coordination nos >8 not recorded correctly by makseq)

   Numbers of cngps with span, min and max seqdif =
       0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000
 2  protein donors in cngp
      0   5   4   1   2   0   0   0   2   0   0    8    6    0    1    5    5    2    0    0    0    0 span
      0   5   4   1   2   0   0   0   2   0   0    8    6    0    1    5    5    2    0    0    0    0 min
 3  protein donors in cngp
      0   0   0   1   0   0   3   1   1   0   0    1    4    3    3    8    2    1    1    2    0    0 span
      0   6  15   3   1   1   0   0   0   0   0    0    3    0    0    1    1    0    0    0    0    0 min
      0   0   1   3   1   0   1   0   0   0   0    1    8    1    4    6    1    2    1    1    0    0 max
 4  protein donors in cngp
      0   0   0   0   0   0   1   1   0   0   0    0    0    0    0    1    1    1    0    1    0    0 span
      0   3   2   1   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
      0   0   1   1   0   0   0   0   0   0   0    0    0    0    0    1    1    1    0    1    0    0 max
 5  protein donors in cngp
      0   0   0   0   0   0   0   0   0   0   0    1    1    0    0    0    0    1    0    0    0    0 span
      0   1   1   1   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
      0   0   0   0   0   1   0   0   0   0   0    1    0    0    0    0    1    0    0    0    0    0 max

   numbers of each type of donor
```

```
     D     E     S     T     O     H     C     K     M     N     Q     R     Y     W     #

    68    37    15    20    51     3     0     1     0    14     3     0     2     0     0


  numbers of donor pairs found with each seqdif

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

      0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000


 DD   0   1   7   0   1   1   0   0   0   0   0    0    1    1    0    2    0    1    0    1    0    0

 DE   0   2   1   0   0   0   0   0   1   0   0    1    2    0    0    0    0    0    0    0    0    0

 DS   0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    1    0    0    0    0    0

 DT   0   0   0   0   1   0   0   0   1   0   0    0    1    0    0    0    0    0    0    0    0    0

 DO   0   2   3   0   0   0   0   0   0   0   0    0    1    0    1    0    0    0    0    0    0    0

 DK   0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    1    0    0    0    0    0    0

 DN   0   3   1   0   0   0   0   0   0   0   0    0    4    0    0    0    1    0    0    0    0    0

 DQ   0   0   0   0   0   0   0   0   0   0   0    1    0    0    0    0    0    0    0    0    0    0

 DY   0   0   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

 ED   0   2   0   0   0   0   0   0   0   0   0    1    3    0    0    2    0    0    0    0    0    0

 EE   0   0   0   0   1   0   0   0   0   0   0    1    1    0    1    0    1    0    0    0    0    0

 EO   0   0   0   0   0   0   0   0   0   0   0    0    0    0    1    2    1    0    0    1    0    0

 EH   0   0   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

 EN   0   0   3   0   0   0   0   0   0   0   0    0    1    0    0    0    0    0    0    0    0    0

 SD   0   0   0   0   0   0   0   0   0   0   0    0    1    0    0    1    0    1    0    0    0    0

 SE   0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    1    0    0    0    0

 SS   0   0   2   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

 ST   0   0   0   0   0   0   0   0   0   0   0    2    1    0    0    1    2    1    0    0    0    0

 TD   0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    1    0    0    0    0    0

 TE   0   0   0   0   1   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

 TS   0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    1    0    0    0    0    0    0

 TT   0   0   0   0   0   0   0   0   0   0   0    3    0    0    0    0    0    0    0    0    0    0

 TO   0   0   1   0   0   0   0   0   0   0   0    0    0    0    0    0    1    0    0    0    0    0

 OD   0   0   0   0   0   1   0   0   0   0   0    0    0    0    0    0    0    1    0    0    0    0

 OE   0   1   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

 OS   0   1   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
```

```
  OT   0   1   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  OO   0   3   1  14   0   0   1   0   0   0   0    1    0    0    0    1    0    0    0    0    0    0

  OH   0   0   1   0   0   0   0   0   0   0   0    0    0    0    0    1    0    0    0    0    0    0

  OQ   0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    1    0    0    0    0    0

  HE   0   0   0   0   0   0   0   0   0   0   0    0    0    0    1    0    0    0    0    0    0    0

  ND   0   0   1   0   0   0   0   0   0   0   0    1    0    0    0    0    0    0    0    0    0    0

  NE   0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    1    0    0    0

  NO   0   0   4   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  QD   0   0   0   0   0   0   0   0   0   0   0    0    0    1    0    0    0    0    0    0    0    0

  QE   0   0   0   0   0   0   0   0   0   0   0    0    1    0    0    0    0    0    0    0    0    0

  QO   0   1   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

  YD   0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    1    0    0    0    0    0    0

  YQ   0   0   0   0   0   0   0   0   0   0   0    0    0    0    1    0    0    0    0    0    0    0


 all   0  17  28  14   4   2   1   0   2   0   0   11   17    2    5   13    9    5    1    2    0    0
```

```
   statistics for              37 Mn coordination groups
   ==================================================


                                   N:    0    1    2    3    4    5    6    7    8
   Numbers with
    total coordn number = N        0    0    0    0    3   13   18    3    0
    N protein donor groups         0    0   11   14   11    1    0    0    0
    N non-protein donors            1    3   15   13    4    1    0    0    0

   (Note: total coord.no includes both Os of bidentate carboxlyate;
   numbers of protein and non protein donors treat a -COO group as one donor.
         also, reminder:  coordination nos >8 not recorded correctly by makseq)

   Numbers of cngps with span, min and max seqdif =
       0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000
2  protein donors in cngp
      0   1   2   1   1   1   0   0   0   0   0    3    0    0    0    0    0    0    2    0    0    0 span
      0   1   2   1   1   1   0   0   0   0   0    3    0    0    0    0    0    0    2    0    0    0 min
3  protein donors in cngp
      0   0   0   0   1   0   0   0   0   0   0    1    0    1    1    5    2    1    2    0    0    0 span
      1   1   3   0   3   1   0   0   1   0   0    1    1    0    1    1    0    0    0    0    0    0 min
      0   0   0   1   0   0   0   0   0   0   0    1    1    2    0    4    3    2    0    0    0    0 max
4  protein donors in cngp
      0   0   0   0   0   0   0   0   0   0   0    1    0    0    1    0    5    1    1    2    0    0 span
      0   1   4   0   2   0   0   0   0   0   0    0    3    1    0    0    0    0    0    0    0    0 min
      0   0   0   0   0   0   0   0   0   1   0    0    0    0    1    4    2    1    1    1    0    0 max
5  protein donors in cngp
      0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    1    0    0    0    0    0 span
      0   1   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
      0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    1    0    0    0    0    0 max

   numbers of each type of donor
```

```
     D     E     S     T     O     H     C     K     M     N     Q     R     Y     W     #

    49    26     3     0     6    22     1     0     0     2     4     0     0     0     0


  numbers of donor pairs found with each seqdif

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

       0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000


 DD    0   0   4   0   1   2   0   0   0   1   0    1    0    1    1    3    2    1    0    0    0    0

 DE    0   0   0   0   1   0   0   0   0   0   0    0    1    1    0    0    1    0    1    0    0    0

 DS    0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    1    0    0    0    0    0    0

 DO    0   3   0   0   0   0   0   0   0   0   0    1    0    0    0    0    0    0    0    0    0    0

 DH    0   1   2   0   2   1   0   0   0   0   0    0    1    0    0    1    1    1    0    0    0    0

 DN    0   0   0   0   0   0   0   0   0   0   0    0    0    1    0    0    0    0    0    0    0    0

 ED    0   0   1   0   0   0   0   0   0   0   0    0    2    0    0    0    0    0    0    0    0    0

 EE    0   0   1   0   0   0   0   0   1   0   0    0    1    0    0    1    0    1    1    0    0    0

 ES    0   0   0   1   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

 EO    0   1   0   0   1   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

 EH    0   0   0   0   0   0   0   0   0   0   0    0    0    0    1    0    1    0    0    0    0    0

 EN    0   0   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

 SD    0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    1    0    0

 OD    0   1   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

 OS    0   0   0   1   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0

 HD    0   0   0   0   1   0   0   0   0   0   0    0    1    1    0    2    1    0    0    0    0    0

 HE    0   0   0   0   0   1   0   0   0   0   0    0    1    1    0    1    0    0    0    0    0    0

 HH    0   0   1   0   0   0   0   0   0   0   0    0    0    1    0    2    0    0    0    0    0    0

 CH    0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    1    0    0    0

 NH    0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    1    0    0    0    0    0    0

 QE    0   0   0   0   0   0   0   0   0   0   0    3    0    0    0    0    0    0    0    0    0    0

 QQ    1   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0


all   1   6  10   2   6   4   0   0   1   1   0    5    7    6    2   12    6    3    3    1    0    0
```

```
   statistics for              38 Fe coordination groups
   =================================================


                                    N:    0    1    2    3    4    5    6    7    8
   Numbers with
    total coordn number = N        0    0    0    0   12   10   16    0    0
    N protein donor groups         0    0    7    9   16    5    1    0    0
    N non-protein donors          10   12    6    5    5    0    0    0    0

   (Note: total coord.no includes both Os of bidentate carboxlyate;
   numbers of protein and non protein donors treat a -COO group as one donor.
         also, reminder:  coordination nos >8 not recorded correctly by makseq)

   Numbers of cngps with span, min and max seqdif =
       0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000
2  protein donors in cngp
      0   0   0   0   0   0   0   0   0   0   0    2    3    0    0    0    0    1    0    1    0    0 span
      0   0   0   0   0   0   0   0   0   0   0    2    3    0    0    0    0    1    0    1    0    0 min
3  protein donors in cngp
      0   0   0   0   1   0   0   0   0   0   0    0    0    1    1    1    1    3    1    0    0    0 span
      0   1   2   1   0   3   0   0   0   0   0    0    0    0    1    1    0    0    0    0    0    0 min
      0   0   0   1   0   0   0   0   0   0   0    0    0    1    1    2    2    2    0    0    0    0 max
4  protein donors in cngp
      0   0   0   0   0   0   0   0   0   0   0    1    1    1    0    5    4    2    0    2    0    0 span
      0   2   2   8   2   0   0   0   0   0   0    0    0    0    0    2    0    0    0    0    0    0 min
      0   0   0   0   0   0   0   0   0   0   0    2    0    2    2    6    1    2    0    1    0    0 max
5  protein donors in cngp
      0   0   0   0   0   0   0   0   0   0   0    0    0    0    1    1    0    0    0    3    0    0 span
      0   0   1   2   1   0   0   0   0   0   0    0    0    1    0    0    0    0    0    0    0    0 min
      0   0   0   0   0   0   0   0   0   0   0    0    1    0    0    1    0    0    1    2    0    0 max
6  protein donors in cngp
      0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    1    0    0    0    0    0 span
```

```
    0  0  1  0  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0   0   0   0 min
    0  0  0  0  0  0  0  0  0  0  0   0   0   0   0   0   1   0   0   0   0   0 max


  numbers of each type of donor
    D    E    S    T    O    H    C    K    M    N    Q    R    Y    W    #
   11   30    0    0    5   60   20    0    3    1    0    0    5    0    1


  numbers of donor pairs found with each seqdif
  -------------------------------------------
     0  1  2  3  4  5  6  7  8  9 10   10- 20- 30- 40- 50-100-150-200-250- 500-   >1000


DD  0  0  0  0  0  0  0  0  0  0  0   0   0   1   0   1   0   0   0   0   0   0
DE  0  0  0  0  0  0  0  0  0  0  0   0   0   1   0   0   0   0   0   0   0   0
DO  0  0  0  0  0  0  0  0  0  0  0   0   0   0   0   0   0   0   1   0   0   0
DH  0  0  0  0  0  0  0  0  0  0  0   0   0   0   0   2   0   0   0   0   0   0
DN  0  0  0  0  0  0  0  0  0  0  0   0   0   1   0   0   0   0   0   0   0   0
DY  0  0  0  1  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0   0   0   0
ED  0  0  0  0  0  0  0  0  0  0  0   0   0   0   1   0   0   0   0   0   0   0
EE  0  0  0  0  0  0  0  0  0  0  0   0   0   6   1   3   0   0   0   0   0   0
EO  0  0  0  0  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0   1   0   0
EH  0  0  1  7  0  0  0  0  0  0  0   1   0   0   0   2   1   0   0   0   0   0
EY  0  0  0  0  0  0  0  0  0  0  0   0   0   0   0   0   1   0   0   0   0   0
HD  0  0  1  0  0  1  0  0  0  0  0   1   0   0   0   0   1   0   0   0   0   0
HE  0  0  2  0  1  0  0  0  0  0  0   0   0   0   2   4   0   1   0   0   0   0
HO  0  0  0  0  0  0  0  0  0  0  0   0   0   0   0   0   0   1   0   0   0   0
HH  0  0  2  0  2  4  2  0  0  0  0   1   5   1   1   2   0   3   0   0   0   0
HC  0  0  0  0  0  0  0  0  0  0  0   0   0   0   0   2   0   0   0   0   0   0
CH  0  0  0  2  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0   0   0   0
CC  0  1  0  6  0  0  0  0  0  0  0   2   1   1   0   0   0   0   0   1   0   0
C#  0  1  0  0  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0   0   0   0
MH  0  0  0  0  0  0  0  0  0  0  0   2   1   0   0   0   0   0   0   0   0   0
NH  0  0  0  0  0  0  0  0  0  0  0   0   0   0   0   1   0   0   0   0   0   0
YH  0  0  0  0  0  0  0  0  0  0  0   1   0   0   0   0   0   1   0   0   0   0
```

```
  YY   0   1   0   0   0   0   0   0   0   0   0    0    0    1    0    0    0    0    0    0    0    0

 all   0   3   6  16   3   5   2   0   0   0   0    8    7   12    5   17    3    6    1    2    0    0
```

```
    statistics for              37 Cu coordination groups
    ==================================================

                               N:   0    1    2    3    4    5    6    7    8
   Numbers with
    total coordn number = N        0    0    0    9   17   11    0    0    0
    N protein donor groups         0    0    1   21   15    0    0    0    0
    N non-protein donors          16   18    3    0    0    0    0    0    0

   (Note: total coord.no includes both Os of bidentate carboxlyate;
   numbers of protein and non protein donors treat a -COO group as one donor.
       also, reminder:  coordination nos >8 not recorded correctly by makseq)

   Numbers of cngps with span, min and max seqdif =
     0  1  2  3  4  5  6  7  8  9 10  10- 20- 30- 40- 50-100-150-200-250- 500-   >1000
2  protein donors in cngp
     0  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    1    0    0 span
     0  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    1    0    0 min
3  protein donors in cngp
     0  1  0  0  0  0  0  0  1  0  0    0    0    4    3    6    0    2    0    4    0    0 span
     1  3  4  1  5  1  1  0  0  1  0    0    0    0    2    2    0    0    0    0    0    0 min
     0  1  0  0  1  0  0  0  0  0  0    0    2    5    1    5    0    2    1    3    0    0 max
4  protein donors in cngp
     0  0  0  0  0  0  0  0  3  0  0    0    0    0    4    7    0    0    0    1    0    0 span
     0  0  5  2  3  5  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0 min
     0  0  0  0  3  0  0  0  0  0  0    0    0    4    3    4    0    0    0    1    0    0 max

   numbers of each type of donor
     D     E     S     T     O     H     C     K     M     N     Q     R     Y     W     #
     1     0     0     0     3    80    26     0    10     0     2     0     1     0     2

   numbers of donor pairs found with each seqdif
```

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

       0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000


  OC  0   0   3   0   0   0   0   0   0   0   0     0     0     0     0     0     0     0     0     0     0     0

  HH  0   3   6   0   3   0   0   0   0   1   0     0     3     2     3     7     0     2     0     5     0     0

  HC  0   0   0   0   0   0   0   0   0   0   0     0     0     7     3     4     0     0     0     0     0     0

  HM  0   0   0   0   0   5   0   0   0   0   0     0     0     0     0     1     0     0     0     0     0     0

  HQ  0   0   0   0   0   2   0   0   0   0   0     0     0     0     0     0     0     0     0     0     0     0

  CO  0   0   3   0   0   0   0   0   0   0   0     0     0     0     0     0     0     0     0     0     0     0

  CH  0   0   0   3   4   6   0   0   0   0   0     0     0     0     0     0     0     0     0     0     0     0

  CC  0   0   0   0   5   0   0   0   0   0   0     0     0     0     0     0     0     0     0     0     0     0

  CM  0   0   0   0   0   0   1   3   0   0   0     0     0     0     0     0     0     0     0     0     0     0

  YH  0   0   0   0   0   0   0   0   0   0   0     0     0     0     0     0     0     0     1     0     0     0

  #D  1   0   0   0   0   0   0   0   0   0   0     0     0     0     0     0     0     0     0     0     0     0

  ##  0   1   0   0   0   0   0   0   0   0   0     0     0     0     0     0     0     0     0     0     0     0


 all  1   4  12   3  12  13   1   3   0   1   0     0     3     9     6    12     0     2     1     5     0     0
```

```
  statistics for            151 Zn coordination groups
  ==================================================

                                  N:    0    1    2    3    4    5    6    7    8
  Numbers with
   total coordn number = N        0    0    7   20   89   32    3    0    0
   N protein donor groups         0    0   22   54   75    0    0    0    0
   N non-protein donors          71   59   19    2    0    0    0    0    0

  (Note: total coord.no includes both Os of bidentate carboxlyate;
  numbers of protein and non protein donors treat a -COO group as one donor.
        also, reminder:  coordination nos >8 not recorded correctly by makseq)

  Numbers of cngps with span, min and max seqdif =
      0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250- 500-    >1000
2  protein donors in cngp
     2   0   5   2   2   0   0   0   0   0   0    1    3    1    0    4    0    1    1    0    0    0 span
     2   0   5   2   2   0   0   0   0   0   0    1    3    1    0    4    0    1    1    0    0    0 min
3  protein donors in cngp
     0   2   0   0   0   0   0   0   0   0   4    3    7    2    8   15    7    2    3    1    0    0 span
     3   2  12   7  16   0   0   2   1   0   1    1    4    3    0    2    0    0    0    0    0    0 min
     0   2   0   0   0   0   3   0   1   1   0    3    8    4    6   14    7    1    3    1    0    0 max
4  protein donors in cngp
     0   0   0   0   0   0   0   0   0   0   0   12   25    6    3   12    9    3    5    0    0    0 span
     4   5  18  36   5   2   0   0   0   0   0    0    1    2    0    2    0    0    0    0    0    0 min
     0   0   0   0   0   1   1   3   2   3   2   20   11    3    4   16    7    1    1    0    0    0 max

  numbers of each type of donor
     D      E      S      T      O      H      C      K      M      N      Q      R      Y      W      #
    61     45      1      0      9    177    196      2      0      1      4      1      0      0      9

  numbers of donor pairs found with each seqdif
```

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

     0  1  2  3  4  5  6  7  8  9 10  10- 20- 30- 40- 50-100-150-200-250- 500-    >1000


DD   0  0  0  1  0  0  0  0  0  0  0    1    1    0    0    1    1    0    0    0    0    0

DE   0  0  2  0  0  0  0  0  0  0  1    1    0    3    0    3    0    1    0    0    0    0

DS   0  0  0  0  0  0  0  0  0  0  0    0    0    0    0    1    0    0    0    0    0    0

DO   0  0  0  0  0  0  0  0  0  0  0    0    0    0    0    2    0    0    0    0    0    0

DH   0  2  2  0  2  0  0  0  0  0  0    3    1    4    0    0    0    0    0    0    0    0

DN   0  0  0  0  0  0  0  0  0  0  0    0    0    1    0    0    0    0    0    0    0    0

DQ   0  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    1    0    0    0    0

ED   0  0  0  0  1  0  0  0  0  0  0    0    0    2    0    2    1    0    1    0    0    0

EE   0  0  0  1  1  0  0  0  0  0  0    0    1    0    0    0    0    0    0    0    0    0

EH   0  0  0  1  0  0  0  0  0  0  0    1    1    1    0    2    4    0    1    0    0    0

EK   0  0  1  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

SD   0  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    1    0    0

OD   1  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

OE   0  0  0  0  0  0  0  0  0  0  0    0    0    0    0    1    0    0    0    0    0    0

OO   0  0  0  0  0  0  0  0  0  0  0    0    0    0    0    1    0    0    0    0    0    0

OH   0  0  0  0  0  1  0  0  0  0  1    1    0    0    0    0    0    0    0    0    0    0

HD   0  1  2  2  3  0  1  2  0  1  0    0    3    1    3    4    3    0    1    0    0    0

HE   0  2  1  2  1  0  0  0  0  0  0    1    4    1    1    3    1    1    0    0    0    0

HH   0  0 11  1 16  0  2  0  1  3  0    2    5    3    5   11    4    0    0    0    0    0

HC   0  0  3  3  1  0  1  1  1  0  0    4    2    1    0    2    0    0    0    0    0    0

CH   0  1  5  4  1  2  0  0  0  0  0    6    5    0    1    4    0    0    0    0    0    0

CC   0  1  9 53  3  9  3  5  3  2  4   12    7    2    2    3    0    0    1    0    0    0

KD   0  0  0  0  0  1  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

NH   0  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    1    0    0    0    0    0

QH   0  0  1  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

QQ   1  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

RH   0  0  0  1  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

#O   5  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

#H   2  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

##   0  2  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0
```

all	9	9	37	69	29	13	7	8	5	6	6	32	30	19	12	40	15	3	4	1	0	0

```
   statistics for              40 Na coordination groups
   ==================================================
                                    N:     0    1    2    3    4    5    6    7    8
   Numbers with
    total coordn number = N          0    0    2    5    5   13   13    2    0
    N protein donor groups           0    0   10   16    7    6    0    1    0
    N non-protein donors            12    9    8    8    3    0    0    0    0

   (Note: total coord.no includes both Os of bidentate carboxlyate;
   numbers of protein and non protein donors treat a -COO group as one donor.
         also, reminder:  coordination nos >8 not recorded correctly by makseq)

   Numbers of cngps with span, min and max seqdif =
       0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000
2  protein donors in cngp
     2   0   2   3   0   0   0   0   0   0   0    0    0    0    1    0    1    1    0    0    0    0 span
     2   0   2   3   0   0   0   0   0   0   0    0    0    0    1    0    1    1    0    0    0    0 min
3  protein donors in cngp
     0   0   0   1   0   2   3   0   0   0   0    0    3    1    1    3    0    1    0    1    0    0 span
     1   4   5   6   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
     0   0   1   5   0   0   0   0   0   0   0    0    3    1    1    3    0    1    0    1    0    0 max
4  protein donors in cngp
     0   0   0   0   0   0   1   0   0   0   0    0    0    0    1    2    2    0    0    1    0    0 span
     2   3   0   0   0   0   0   0   0   0   0    0    0    0    2    0    0    0    0    0    0    0 min
     0   0   0   1   0   0   0   0   0   0   0    0    0    0    2    3    0    0    0    1    0    0 max
5  protein donors in cngp
     0   0   0   0   0   0   0   2   0   0   0    1    1    1    1    0    0    0    0    0    0    0 span
     2   3   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
     0   0   0   2   0   0   0   0   1   0   0    0    3    0    0    0    0    0    0    0    0    0 max
7  protein donors in cngp
     0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    1    0    0    0    0    0 span
     1   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
```

```
      0  0  0  0  0  0  0  0  0  0  0   0    0    0    0    0    1    0    0    0    0    0 max


    numbers of each type of donor

      D    E    S    T    O    H    C    K    M    N    Q    R    Y    W    #

     16   11    4    2   92    0    0    1    0    6    1    0    0    0    0


    numbers of donor pairs found with each seqdif

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

       0  1  2  3  4  5  6  7  8  9 10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000


  DD  0  0  0  0  0  0  1  0  0  0  0    1    1    0    0    0    0    0    0    0    0    0

  DE  0  0  0  0  0  0  0  0  0  0  0    0    1    0    0    0    0    0    0    0    0    0

  DS  0  0  1  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

  DO  0  1  2  0  0  0  0  0  0  0  0    0    0    0    0    1    0    0    0    1    0    0

  DN  0  0  0  1  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

  EE  0  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    1    0    0    0    0

  EO  0  2  0  1  0  1  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

  SD  0  0  1  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

  SO  0  2  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

  TO  0  1  0  0  0  0  0  0  0  0  0    0    0    0    0    0    1    0    0    0    0    0

  OD  1  1  2  1  0  0  0  0  1  0  0    0    0    0    0    0    0    0    0    0    0    0

  OE  3  1  0  1  0  0  0  0  0  0  0    0    0    0    2    0    0    0    0    0    0    0

  OS  1  0  0  0  0  0  0  0  0  0  0    0    1    0    0    0    0    0    0    0    0    0

  OT  2  0  0  0  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

  OO  0  4 12 15  0  0  0  0  0  0  0    0    3    1    5    5    0    1    0    1    0    0

  ON  1  0  0  2  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

  OQ  0  0  0  1  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

  KE  0  0  0  0  0  0  0  0  0  0  0    0    0    0    1    0    0    0    0    0    0    0

  ND  0  0  0  1  0  0  0  0  0  0  0    0    0    0    0    0    0    0    0    0    0    0

  NO  0  1  0  0  0  0  0  0  0  0  0    0    0    0    0    0    1    0    0    0    0    0


 all  8 13 18 23  0  1  1  0  1  0  0    1    6    1    8    6    2    2    0    2    0    0
```

```
    statistics for              28 K coordination groups
    ================================================

                                  N:    0    1    2    3    4    5    6    7    8
   Numbers with
    total coordn number = N        0    0    0    2    5    5    6    8    2
    N protein donor groups         0    0    8    4    6    7    2    1    0
    N non-protein donors           5    5   11    5    0    2    0    0    0

   (Note: total coord.no includes both Os of bidentate carboxlyate;
   numbers of protein and non protein donors treat a -COO group as one donor.
         also, reminder:  coordination nos >8 not recorded correctly by makseq)

   Numbers of cngps with span, min and max seqdif =
      0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000
2  protein donors in cngp
     1   3   1   0   0   1   0   0   0   0   0    1    0    0    0    1    0    0    0    0    0    0 span
     1   3   1   0   0   1   0   0   0   0   0    1    0    0    0    1    0    0    0    0    0    0 min
3  protein donors in cngp
     0   1   0   0   1   0   0   0   0   0   0    0    0    2    0    0    0    0    0    0    0    0 span
     1   0   3   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
     0   1   1   0   0   0   0   0   0   0   0    0    2    0    0    0    0    0    0    0    0    0 max
4  protein donors in cngp
     0   0   0   0   0   0   2   0   0   0   0    1    0    1    0    1    0    1    0    0    0    0 span
     0   5   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
     0   0   0   2   0   0   0   0   0   0   0    1    0    2    0    0    0    1    0    0    0    0 max
5  protein donors in cngp
     0   0   0   0   0   0   0   1   0   0   0    0    0    1    0    2    0    1    2    0    0    0 span
     3   2   1   0   0   0   0   0   0   0   0    0    1    0    0    0    0    0    0    0    0    0 min
     0   0   0   0   0   1   0   0   0   0   0    0    2    1    0    1    0    1    1    0    0    0 max
6  protein donors in cngp
     0   0   0   0   0   0   0   0   0   0   0    0    0    1    0    0    0    1    0    0    0    0 span
```

```
      0   1   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
      0   0   0   0   0   0   0   0   0   0   0    0    1    0    0    0    1    0    0    0    0    0 max
 7   protein  donors  in  cngp
      0   0   0   0   0   0   0   0   0   0   0    0    1    0    0    0    0    0    0    0    0    0 span
      1   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0 min
      0   0   0   0   0   0   0   0   0   0   0    1    0    0    0    0    0    0    0    0    0    0 max


    numbers  of  each  type  of  donor
      D     E     S     T     O     H     C     K     M     N     Q     R     Y     W     #
      7    13     7     9    57     0     0     0     0     9     4     0     0     0     0


    numbers  of  donor  pairs  found  with  each  seqdif
    --------------------------------------------
        0   1   2   3   4   5   6   7   8   9  10   10-  20-  30-  40-  50-100-150-200-250-  500-    >1000


   DO   0   0   0   0   0   2   0   0   0   0   0    0    0    0    0    1    0    0    0    0    0    0
   EE   0   0   0   0   0   0   0   0   0   0   0    1    0    0    0    1    0    2    0    0    0    0
   ET   0   0   0   0   0   0   0   0   0   0   0    0    2    0    0    0    0    0    0    0    0    0
   EO   0   2   0   0   0   0   0   0   0   0   0    0    1    0    0    0    0    0    0    0    0    0
   EN   0   0   1   0   0   0   0   0   0   0   0    0    1    0    0    0    0    0    0    0    0    0
   EQ   0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    1    0    0    0    0    0    0
   SD   0   0   0   0   0   0   0   0   0   0   0    0    0    0    0    0    1    0    0    0    0    0
   SE   0   1   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
   SS   0   0   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
   SO   0   0   0   0   0   0   0   0   0   0   0    0    1    0    0    0    0    0    1    0    0    0
   TD   0   0   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
   TT   0   0   0   0   0   0   0   0   0   0   0    1    0    0    0    0    0    0    0    0    0    0
   TO   0   1   1   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
   TN   0   0   0   0   0   0   0   0   0   0   0    0    0    2    0    0    0    0    0    0    0    0
   OD   1   0   3   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
   OE   0   2   0   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
   OS   1   0   2   1   0   2   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
   OT   4   0   2   0   0   0   0   0   0   0   0    0    0    0    0    0    0    0    0    0    0    0
```

OO	0	8	5	5	0	0	0	0	0	0	0	1	2	1	0	1	0	0	0	0	0	0
ON	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NO	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
NN	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NQ	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
QE	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
QO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
all	8	16	19	7	0	4	0	0	0	0	0	3	9	3	0	5	1	2	1	0	0	0

ZN & CA COORDINATION GROUPS IN PROTEINS which occur more than once in 30% cull set

The data includes, for the individual occurrences:

i) amino-acid sequences, starting at *relseq*=-10; *relseq*=0 is the first donor amino-acid

ii) sequences of amino-acid types - g is glycine, s is small hydrophobic, h is large hydrophobic, a is acidic, b is basic, p is polar

iii) sequences of (Efimov type) conformations

iv) starting residue number, resolution, PDB header name,

v) coordination group definition (as in Table 3D)

vi) summary of analysis of agreement of torsion angles over selected regions of protein chain; chains are identified by their protein chain numbers (pc) - this is the first number in each row of the tables.

Ca EOOE 2 3 5

Ca QDONOO 39 34 2 2 2

Ca DDDOE,DNDOE,DDNOE 2 2 2 5

Ca DDDOD 2225 and related

Ca DDOO 2 2 2

.

For Zn coordination groups the data included is like that for Ca, except that details are also given of individual examples of these coordination groups when they occur in proteins in the 90% cull set, and the conformations are described by the categories of Hovmoller et al., using a for helix, b for sheet, g for turn, and j for other.

Zn CCCC 3 3 8

Zn CCCC 3 17 3

Zn HHH 4 6

```
Ca   EOOE  2  3  5   and  related


    amino-acid  sequences  ****

     start  of  selected  part  of  cngroup          |

     relseq                 -10          -5           0           5          10          15

  1   1sbw  from    60  A:   K  S  G  I  Q  V  R  _  _  _  E  D  N  I  N  V  V  E  G  N  E  Q  F  I  S  A  S  K  S  I  V

  2   2btc  from    60  E:   K  S  G  I  Q  V  R  _  _  _  E  D  N  I  N  V  V  E  G  N  E  Q  F  I  S  A  S  K  S  I  V

  3   1avw  from    60  A:   K  S  R  I  Q  V  R  _  _  _  E  H  N  I  D  V  L  E  G  N  E  Q  F  I  N  A  A  K  I  I  T

  4   1pyt  from   460  C:   S  T  R  T  Y  Q  V  V  L  G  E  Y  D  R  S  V  L  Q  G  S  E  Q  V  I  P  I  G  D  L  F  V

                                                          *  see  below

    aa  types  -  acidic,  etc  **

     relseq                 -10          -5           0           5          10          15

  1   1sbw  from    60  A:   b  s  g  h  p  h  b  _  _  _  a  a  p  h  p  h  h  a  g  p  a  p  h  h  s  s  s  b  s  h  h

  2   2btc  from    60  E:   b  s  g  h  p  h  b  _  _  _  a  a  p  h  p  h  h  a  g  p  a  p  h  h  s  s  s  b  s  h  h

  3   1avw  from    60  A:   b  s  b  h  p  h  b  _  _  _  a  p  p  h  a  h  h  a  g  p  a  p  h  h  p  s  s  b  h  h  s

  4   1pyt  from   460  C:   s  s  b  s  h  p  h  h  h  g  a  h  a  b  s  h  h  p  g  s  a  p  h  h  h  h  g  a  h  h  h


    conformation  sequences  **

     relseq                 -10          -5           0           5          10          15

  1   1sbw  from    60  A:   .  b  g  b  b  b  b  .  .  .  b  .  b  k  k  b  b  k  .  g  b  b  b  b  b  b  a  b  b  b  b

  2   2btc  from    60  E:   b  b  g  b  b  b  b  .  .  .  b  a  b  k  k  b  b  k  g  k  b  b  b  b  b  b  a  b  b  b  b

  3   1avw  from    60  A:   b  a  b  b  b  b  b  .  .  .  b  a  b  k  k  b  b  k  g  k  b  b  b  b  b  b  a  b  b  b  b

  4   1pyt  from   460  C:   b  .  .  b  b  b  b  b  b  g  b  .  b  k  k  b  b  k  g  a  b  b  b  b  b  b  k  k  b  d  b


  **********   protein  names  (pdb  header)  *******

  1   1sbw  at    70  A    1.80       HYDROLASE/HYDROLASE  INHIBITOR  Trypsin-like  serine  proteases

  2   2btc  at    70  E    1.50       HYDROLASE/HYDROLASE  INHIBITOR  Trypsin-like  serine  proteases

  3   1avw  at    70  A    1.75       COMPLEX  (PROTEINASE/INHIBITOR)  Trypsin-like  serine  proteases

  4   1pyt  at   470  C    2.35       TERNARY  COMPLEX  (ZYMOGEN)        Trypsin-like  serine  proteases


  **********   cngps   ************

  1  1   EOOE       Ca    2    3    5   -1   -1   -1   -1  ....   6  1sbw   0  0.2  1.8   CA   801  GLUA   70  ....     |    ZZ   3.
```

```
 2 1  EOOE     Ca   2   3   5  -1  -1  -1  -1 ....  6 2btc  0 0.2 1.5  CA  700 GLUE  70 ....    |   ZZ  3.
 3 1  EOOEE    Ca   2   3   2   3  -1  -1  -1 ..... 6 1avw  0 0.1 1.8  CA  700 GLUA  70 .....   |    Z  3.
 4 1  EOOQE    Ca   2   3   2   3  -1  -1  -1 ..... 5 1pyt  1 0.3 2.3  CA  650 GLUC 470 .....   |       3.

 * aa sequences:  in first three entries, there is no residue number 68,
                       residues 67 and 69 are L G, and numbering anomaly seems
                       to have thrown my program tortab  (found 2.7.02, tortab12a now corrected))




               3 protein chains to be compared
start and end seq values for fi   1  10; for psi    0   9

  relseq     meanfi          meanpsi      number
     0        -89( 17)        146(  3)          3        b
     1       -131(  7)        -66( 10)          3        .     near a
     2        -91( 14)        107(  3)          3        b
     3        -70(  5)          -6(  3)          3        k
     4       -101(  5)         -7( 11)          3        k
     5       -129( 17)        132( 10)          3        b
     6        -72(  9)        124( 10)          3        b
     7       -114( 14)         -7(  8)          3        k
     8        101(  7)        -13(  7)          3        g
     9        -99(  7)          -3( 18)          3        k
    10        -74( 13)        148(  3)          3        b

1 is like them over relseq 1-7, different at relseq 9-10

                4 protein chains to be compared
start and end seq values for fi   1  10; for psi    0   9

  relseq     meanfi          meanpsi      number
     0        -89( 14)        147(  3)          4        b
```

```
     1       -131(  6)        -69( 10)           4         .      near a
     2         -88( 12)        108(  3)           4         b
     3         -71(  4)          -7(  3)           4         k
     4       -101(  4)          -6(  9)           4         k
     5       -129( 14)        133(  9)           4         b
     6         -75(  9)        121( 10)           4         b
     7       -114( 11)          -8(  6)           4         k
     8         104(  7)          26( 81)           4         g
     9         126(166)          -3( 15)           4         .
    10         -75( 11)        148(  3)           4         b

all belong to same fold family

local conformations same from relseq -7 to relseq + 7,
local conformations same from relseq -7 to relseq +14 for pcs 2 3 4



 1  1sbw at   70 A   1.80  Ca 801
        4 donors, 2 wat, CN 6,  delta oct 13 deg, six good dists

 2  2btc at   70 E   1.50  Ca 700
        4 donors, 2 wat, CN 6,  delta oct 11 deg, six fairly good dists

 3  1avw at   70 A   1.75  Ca 700
        5 donors, 1 wat, CN 6,  delta oct 9 deg, six good dists

 4  1pyt at  470 C   2.35  Ca 650
        5 donors, CN 5, angles 68-94, and 159(two) - if oct  v distorted
          donor 4 (Q) is at 2.91 A, other dists look a bit odd
```

```
cagp2:    QDONOO    Ca   39   34    2    2    2


has been done in parts

a) QD   39

   amino-acid sequences ****

    start of selected part of cngroup         |

    relseq                -10         -5          0          5         10         15

 1   1cse from    -8 E:  _ _ _ _ _ _ _ _ _ A Q T V P Y G I P L I K A D K V Q A Q G F K G A N V K V A V L D T G I Q A S H P

 2   1gci from    -8  :  _ _ _ _ _ _ _ _ _ A Q S V P W G I S R V Q A P A A H N R G L T G S G V K V A V L D T G I S _ T H P

 3   1scj from    -8 A:  _ _ _ _ _ _ _ _ _ A Q S V P Y G I S Q I K A P A L H S Q G Y T G S N V K V A V I D S G I D S S H P

 4   2sic from    -8 E:  _ _ _ _ _ _ _ _ _ A Q S V P Y G V S Q I K A P A L H S Q G Y T G S N V K V A V I D S G I D S S H P


   aa types - acidic, etc **

    relseq                -10         -5          0          5         10         15

 1   1cse from    -8 E:  _ _ _ _ _ _ _ _ _ s p s h h h g h h h h b s a b h p s p g h b g s p h b h s h h a s g h p s s p h

 2   1gci from    -8  :  _ _ _ _ _ _ _ _ _ s p s h h h g h s b h p s h s s p p b g h s g s g h b h s h h a s g h s _ s p h

 3   1scj from    -8 A:  _ _ _ _ _ _ _ _ _ s p s h h h g h s p h b s h s h p s p g h s g s p h b h s h h a s g h a s s p h

 4   2sic from    -8 E:  _ _ _ _ _ _ _ _ _ s p s h h h g h s p h b s h s h p s p g h s g s p h b h s h h a s g h a s s p h


   conformation sequences **

    relseq                -10         -5          0          5         10         15

 1   1cse from    -8 E:  . . . . . . . . . . b b b b a a a a a k g a a a a a a a g b k g b g b b b b b b . k j b b a k b a

 2   1gci from    -8  :  . . . . . . . . . . b b b b a a a a a k g a a a a a a k g b k g b g b b b b b b . k j b . . . b k

 3   1scj from    -8 A:  . . . . . . . . . . b b b b a a a a a k g a a a a a a k g b k g b g b b b b b b . k j b b k k b k

 4   2sic from    -8 E:  . . . . . . . . . . b b b b a a a a a k g a a a a a a k g b k g b g b b b b b b . k j b b k k b a



 **********   cngps   ************

 1 1   QDONOO    Ca   39   34    2    2    2   -1   -1 ......6 1cse   0 0.1 1.2   CA   430 GLNE    2 .b....   |         3.

 2 1   QDONOO    Ca   39   34    2    2    2   -1   -1 ......6 1gci   0 0.1 0.8   CA   277 GLN     2 .b....   |         3.

 3 1   QDONOO    Ca   39   34    2    2    2   -1   -1 ......6 1scj   0 0.2 2.0   CA   381 GLNA    2 .b....   |         3.

 4 1   QDONOO    Ca   39   34    2    2    2   -1   -1 ......6 2sic   0 0.2 1.8   CA   501 GLNE    2 .b....   |         3.
```

```
all torsion angles along this part of chain have s.d. <= 13 deg, except
    that there is one residue missing, at relseq 35 in 1gci



b) DONO 34 2 2

  amino-acid sequences ****
   start of selected part of cngroup         |
   relseq                -10          -5           0            5          10          15
 1  1cse from   31 E:  L D T G I Q A S H P D L N V V G G A S F V A G E A _ Y N T D G N G H G T H V A G T V A A L D N T T
 2  1gci from   31  :  L D T G I S _ T H P D L N I R G G A S F V P G E P S T _ Q D G N G H G T H V A G T I A A L N N S I
 3  1scj from   31 A:  I D S G I D S S H P D L N V R G G A S F V P S E T N P Y Q D G S S H G T H V A G T I A A L N N S I
 4  2sic from   31 E:  I D S G I D S S H P D L K V A G G A S M V P S E T N P F Q D N N S H G T H V A G T V A A L N N S I

  aa types - acidic, etc **
   relseq                -10          -5           0            5          10          15
 1  1cse from   31 E:  h a s g h p s s p h a h p h h g g s s h h s g a s _ h p s a g p g p g s p h s g s h s s h a p s s
 2  1gci from   31  :  h a s g h s _ s p h a h p h b g g s s h h h g a h s s _ p a g p g p g s p h s g s h s s h p p s h
 3  1scj from   31 A:  h a s g h a s s p h a h p h b g g s s h h h s a s p h h p a g s s p g s p h s g s h s s h p p s h
 4  2sic from   31 E:  h a s g h a s s p h a h b h s g g s s h h h s a s p h h p a p p s p g s p h s g s h s s h p p s h

  conformation sequences **
   relseq                -10          -5           0            5          10          15
 1  1cse from   31 E:  b . k j b b a k b a k b b b k . . b b k b b g b . . . k b b a a g a a a a a a a a a k b b b . k b
 2  1gci from   31  :  b . k j b . . . b k k b b b k . . . b k b b g d a b . . . b k a . a a a a a a a a a . b b b . k b
 3  1scj from   31 A:  b . k j b b k k b k k b b b k . . . b k b a k d k b k k b b k a g a a a a a a a a a d b b b . k b
 4  2sic from   31 E:  b . k j b b k k b a k b b b a . . b b k b a k b k b a k b b k a g a a a a a a a a a d b b b . d b



 **********  cngps  ************
 1 2  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1cse  0 0.1 1.2  CA  430 GLNE   2 .b....  |        3.
 2 2  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1gci  0 0.1 0.8  CA  277 GLN    2 .b....  |        3.
```

```
  3 2   QDONOO    Ca   39   34    2    2    2   -1   -1 ......6 1scj   0 0.2 2.0   CA   381 GLNA    2 .b....   |         3.
  4 2   QDONOO    Ca   39   34    2    2    2   -1   -1 ......6 2sic   0 0.2 1.8   CA   501 GLNE    2 .b....   |         3.

note missing residue numbers in 34 loop in 1 and 2, but residue numbering has been done
  so that the rest aligns well
torsion angles agree well over whole range, with s.d. <= 15 deg,EXCEPT
  i) small increases to 20-30 deg around missing residues
  ii) major anomaly at relseq 12:  in pc 1 2 this is G with conformation g
                                        in pc 3 4 this is S with conformation k



c) ONOO 2 2 2 part of QDONOOO

  amino-acid sequences ****
    start of selected part of cngroup       |
    relseq               -10        -5          0           5        10           15
  1  1cse from    65 E:  G T H V A G T V A A L D N T T G V L G V A P S V S L
  2  1gci from    65  :  G T H V A G T I A A L N N S I G V L G V A P S A E L
  3  1scj from    65 A:  G T H V A G T I A A L N N S I G V L G V S P S A S L
  4  2sic from    65 E:  G T H V A G T V A A L N N S I G V L G V A P S A S L

  aa types - acidic, etc **
    relseq               -10        -5          0           5        10           15
  1  1cse from    65 E:  g s p h s g s h s s h a p s s g h h g h s h s h s h
  2  1gci from    65  :  g s p h s g s h s s h p p s h g h h g h s h s s a h
  3  1scj from    65 A:  g s p h s g s h s s h p p s h g h h g h s h s s s h
  4  2sic from    65 E:  g s p h s g s h s s h p p s h g h h g h s h s s s h

  conformation sequences **
    relseq               -10        -5          0           5        10           15
  1  1cse from    65 E:  a a a a a a a a k b b b . k b j b b b a b k k b b b
  2  1gci from    65  :  a a a a a a a a . b b b . k b j b b b a d k k b b b
```

```
 3  1scj from   65 A:  a a a a a a a a d b b b . k b . b b b a b k k b b b
 4  2sic from   65 E:  a a a a a a a a d b b b . d b j b b b a d a k b b b


 **********  protein names (pdb header) *******
 1  1cse at   75 E   1.2       COMPLEX(SERINE PROTEINASE-INHIBITOR) Subtilisin-like
                                        Subtilisin from Bacillus subtilis, carlsberg
 2  1gci at   75      0.78      SUBTILISIN FROM BACILLUS LENTUS     Subtilisin-like
                                        Subtilisin from Bacillus lentus
 3  1scj at   75 A   2.0       HYDROLASE                           Subtilisin-like
                                        Subtilisin from Bacillus subtilis, E
 4  2sic at   75 E   1.8       COMPLEX (PROTEINASE/INHIBITOR)  Subtilisin-like
                                        Subtilisin from Bacillus amyloliquefaciens, Novo/BPN'


 **********  cngps  ************
 1 3  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1cse  0 0.1 1.2  CA  430 GLNE   2 .b....  |        3.
 2 3  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1gci  0 0.1 0.8  CA  277 GLN    2 .b....  |        3.
 3 3  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1scj  0 0.2 2.0  CA  381 GLNA   2 .b....  |        3.
 4 3  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 2sic  0 0.2 1.8  CA  501 GLNE   2 .b....  |        3.

chains agree over whole range within 8 deg

 relseq     meanfi           meanpsi      number
     0        -64(  4)        149(  4)         4        b
     1        -90(  8)        102(  2)         4        b
     2       -153(  2)       -149(  4)         4        .   near a
     3       -108(  7)         19(  9)         4        k
     4       -135(  7)        164(  2)         4        b
     5         62(  5)       -131(  3)         4        j
     6       -118(  2)       -165(  4)         4        b



 1  1cse at   75 E   1.2
        D is bidentate, CN 7, O...O dists 3.0-4.9 (except bidentate D)
```

```
2   1gci at    75       0.78

        D is bidentate,  CN 7, O...O dists 3.0-4.7 (except bidentate O)


3   1scj at    75 A    2.0

        D is bidentate,  CN 7, O...O dists 3.1-5.3 (except bidentate O)


4   2sic at    75 E    1.8

        D is bidentate,  CN 7, O...O dists 3.0-4.7 (except bidentate O)
```

```
Ca      DDDOE or DDNOE or DNDOE    2 2 2 5



   amino-acid sequences ****

    start of selected part of cngroup      |

    relseq               -10        -5         0          5         10        15

  1  1cdl from    10 A:  A E F K E A F S L F D K D G D G T I T T K E L G T V M R S L G

  2  1cdl from   119 B:  E E V D E M I R E A D I D G D G Q V N Y E E F V Q M M T _ _ _

  3  1acc from   167  :  _ _ _ _ _ _ _ _ V P D R D N D G I P D S L E V E G Y T V D V K

  4  1g4y from    10 R:  A E F K E A F S L F D K D G D G T I T T K E L G T V M R S L G

  5  1sra from   247  :  H C T T R F F E T C D L D N D K Y I A L D E W A G C F G I K Q

  6  1vrk from    10 A:  A E F K E A F S L F D K D G D G T I T T K E L G T V M R S L G

  7  1vrk from   119 A:  E E V D E M I R E A D V D G D G Q V N Y E E F V Q V M M A K _

  8  2pvb from    80 A:  A E T K A F L A D G D K D G D G M I G V D E F A A M I K A _ _

  9  1a2x from   129 A:  E E I E S L M K D G D K N N D G R I D F D E F L K M M E G V Q

 10  1a2x from    93 A:  E E L A E C F R I F D R N A D G Y I D A E E L A E I F R A S G

 11  2sas from    60  :  D E W R D L K G R A D I N K D D V V S W E E Y L A M W E K T I

 12  2scp from    94 A:  G P L P L F F R A V D T N E D N N I S R D E Y G I F F G M L G

 13  2scp from   128 A:  T M A P A S F D A I D T N N D G L L S L E E F V I A G S D F F

 14  1cdl from    83 A:  E E I R E A F R V F D K D G N G Y I S A A E L R H V M T N L G

 15  1cdl from    46 A:  A E L Q D M I N E V D A D G N G T I D F P E F L T M M A R K M

 16  1g4y from    46 R:  A E L Q D M I N E V D A D G N G T I D F P E F L T M M A R K M

 17  1rec from   100  :  Q K L E W A F S L Y D V D G N G T I S K N E V L E I V T A I F

 18  1vrk from    46 A:  A E L Q D M I N E V D A D G N G T I D F P E F L N L M A R K M


   aa types - acidic, etc **

    relseq               -10        -5         0          5         10        15

  1  1cdl from    10 A:  s a h b a s h s h h a b a g a g s h s s b a h g s h h b s h g

  2  1cdl from   119 B:  a a h a a h h b a s a h a g a g p h p h a a h h p h h s _ _ _

  3  1acc from   167  :  _ _ _ _ _ _ _ _ h h a b a p a g h h a s h a h a g h s h a h b

  4  1g4y from    10 R:  s a h b a s h s h h a b a g a g s h s s b a h g s h h b s h g

  5  1sra from   247  :  p h s s b h h a s h a h a p a b h h s h a a h s g h h g h b p
```

```
  6   1vrk  from    10  A:   s a h b a s h s h h a b a g a g s h s s b a h g s h h b s h g
  7   1vrk  from   119  A:   a a h a a h h b a s a h a g a g p h p h a a h h p h h h s b _
  8   2pvb  from    80  A:   s a s b s h h s a g a b a g a g h h g h a a h s s h h b s _ _
  9   1a2x  from   129  A:   a a h a s h h b a g a b p p a g b h a h a a h h b h h a g h p
 10   1a2x  from    93  A:   a a h s a h h b h h a b p s a g h h a s a a h s a h h b s s g
 11   2sas  from    60   :   a a h b a h b g b s a h p b a a h h s h a a h h s h h a b s h
 12   2scp  from    94  A:   g h h h h h h b s h a s p a a p p h s b a a h g h h h g h h g
 13   2scp  from   128  A:   s h s h s s h a s h a s p p a g h h s h a a h h h s g s a h h
 14   1cdl  from    83  A:   a a h b a s h b h h a b a g p g h h s s s a h b p h h s p h g
 15   1cdl  from    46  A:   s a h p a h h p a h a s a g p g s h a h h a h h s h h s b b h
 16   1g4y  from    46  R:   s a h p a h h p a h a s a g p g s h a h h a h h s h h s b b h
 17   1rec  from   100   :   p b h a h s h s h h a h a g p g s h s b p a h h a h h s s h h
 18   1vrk  from    46  A:   s a h p a h h p a h a s a g p g s h a h h a h h p h h s b b h


    conformation  sequences  **

     relseq                  -10         -5          0           5          10          15

  1   1cdl  from    10  A:   a a a a a a a a a a d a k g k g b b b a a a a a a a a a a d g
  2   1cdl  from   119  B:   a a a a a a a a a k b a k g k g b b b a a a a a k a k . . . .
  3   1acc  from   167   :   . . . . . . . . . b d a k g k g b b a a a a a a . b b b b b b
  4   1g4y  from    10  R:   a a a a a a a a k a b a k g k g b b b a a a a a a a a a k k g
  5   1sra  from   247   :   k k a a a a a a a k d a k g k g b b b a a a a a a a k g b b a
  6   1vrk  from    10  A:   a a a a a a a a a a d a k g k g b b b a a a a a a a a a k k g
  7   1vrk  from   119  A:   a a a a a a a a a k b a k g k g b b b a a a a a a a a k b . .
  8   2pvb  from    80  A:   a a a a a a a a a a d k k g k g b b b a a a a a a a a k . . .
  9   1a2x  from   129  A:   a a a a a a a a a a b a k g k g b b b a a a a a a a k b g b .
 10   1a2x  from    93  A:   a a a a a a a a a a b a k g k g b b b a a a a a a a k a k b k
 11   2sas  from    60   :   a a a a a a a k a a b a k g k g b b b a a a a a a a a a a a a
 12   2scp  from    94  A:   a k a a a a a a a a d a k g k g b b b a a a a a a a a a a k g
 13   2scp  from   128  A:   k k a a a a a a a a d a k g k g b b b a a a a a a a a a a a a
 14   1cdl  from    83  A:   a a a a a a a k a a d a k g k g b b b a a a a a a a a a a k g
 15   1cdl  from    46  A:   a a a a a a a a a k b a k g k g b b b a k a a a a a a k b b b
 16   1g4y  from    46  R:   a a a a a a a a a d b a k g k g b b b a a a k a a a a a k d a
 17   1rec  from   100   :   k k a a a a a a a a d a k g k g b b b a a a a a a a a a a a a
```

```
18  1vrk from   46 A:  a a a a a a a a a k d a k g k g b b b a a a a a a a a a b b b


  **********  protein names (pdb header) *******

  1  1cdl at   20 A   2.2       CALCIUM-BINDING PROTEIN      EF Hand-like Calmodulin-like

  2  1cdl at  129 B   2.2       CALCIUM-BINDING PROTEIN

  3  1acc at  177     2.1       TOXIN     Anthrax protective antigen

  4  1g4y at   20 R   1.6       SIGNALING PROTEIN           EF Hand-like Calmodulin-like

  5  1sra at  257     2.0       CALCIUM-BINDING PROTEIN          EF Hand-like Osteonectin

  6  1vrk at   20 A   1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT   Calmodulin (fold)

  7  1vrk at  129 A   1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT

  8  2pvb at   90 A   0.91      METAL BINDING PROTEIN     EF Hand-like   Parvalbumin

  9  1a2x at  139 A   2.3       COMPLEX (SKELETAL MUSCLE/MUSCLE PROT   EF Hand-like Calmodulin-like

 10  1a2x at  103 A   2.3       COMPLEX (SKELETAL MUSCLE/MUSCLE PROT

 11  2sas at   70     2.4       CALCIUM-BINDING PROTEIN   EF Hand-like Calmodulin-like

 12  2scp at  104 A   2.0       BINDING PROTEIN     EF Hand-like Calmodulin-like

 13  2scp at  138 A   2.0       BINDING PROTEIN

 14  1cdl at   93 A   2.2       CALCIUM-BINDING PROTEIN

 15  1cdl at   56 A   2.2       CALCIUM-BINDING PROTEIN

 16  1g4y at   56 R   1.6       SIGNALING PROTEIN

 17  1rec at  110     1.9       CALCIUM-BINDING PROTEIN   EF Hand-like  Calmodulin-like

 18  1vrk at   56 A   1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT


  **********  cngps  ************

  1 1  DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  0 0.4 2.2  CA    1 ASPA  20 .b..b   |

  2 1  DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  0 0.4 2.2  CA    4 ASPB 129 .b..b   |

  3 1  DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1acc  1 0.2 2.1  CA  800 ASP  177 .....   |     Z

  4 1  DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1g4y  0 0.2 1.6  CA 1001 ASPR  20 ....b   |     Z

  5 1  DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1sra  0 0.2 2.0  CA  302 ASP  257 ....b   |     Z

  6 1  DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.1 1.9  CA  151 ASPA  20 ....b   |     Z  2.

  7 1  DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.2 1.9  CA  154 ASPA 129 ....b   |     Z  2.

  8 1  DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 2pvb  0 0.1 0.9  CA  111 ASPA  90 ....b   |     Z

  9 1  DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1a2x  1 0.4 2.3  CA  161 ASPA 139 ..b.b   |

 10 1  DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1a2x  2 0.2 2.3  CA  160 ASPA 103 ....b   |
```

```
11 1  DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 2sas  0 0.4 2.4  CA  187 ASP   70 ....b   |     Z
12 1  DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 2scp  0 0.1 2.0  CA  191 ASPA 104 ....b   |     Z
13 1  DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 2scp  0 0.2 2.0  CA  192 ASPA 138 ....b   |     Z
14 1  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  0 0.3 2.2  CA    3 ASPA  93 .b..b   |
15 1  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  1 0.4 2.2  CA    2 ASPA  56 ....b   |
16 1  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1g4y  0 0.4 1.6  CA 1002 ASPR  56 ....b   |
17 1  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1rec  0 0.2 1.9  CA  501 ASP  110 ....b   |     Z
18 1  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.1 1.9  CA  152 ASPA  56 ....b   |     Z  2.
```

excluding 3 (1acc) agreement over whole cngp very close (relseq 1-11, 0-10)

```
  relseq     meanfi          meanpsi      number

     0        -78(  7)         84( 12)        17        b
     1        -65( 14)        -40( 16)        17        a
     2        -86( 12)          2( 11)        17        k
     3         58( 10)         28( 12)        17        g
     4        -86( 12)          5(  8)        17        k
     5         81( 12)          3( 12)        17        g
     6       -134(  9)        159( 16)        17        b
     7       -109( 12)        121(  5)        17        b
     8        -94(  9)        172(  9)        17        b
     9        -55( 11)        -44( 10)        17        a
    10        -60(  7)        -39(  8)        17        a
    11        -68(  6)        -36(  5)        17        a
```

before cngp, i.e. relseq -10 to 0, conformations all same

after cngp, some diversity: from relseq 17 onwards

1 4 6 12 14 are similar

7 8 15 18 are similar

11 13 17 are similar

```
Ca DDDOD etc   all with 2 2 2 5 or related sequence


   and, for comparison,

       protein chains 1 and 2 are the models from the set of 18 DDDOE 2225 groups


  amino-acid sequences ****

   start of selected part of cngroup        |

   relseq             -10         -5          0          5         10         15

 1  1acc from  167   :  _ _ _ _ _ _ _ _ V P D R D N D G I P D S L E V E G Y T V D V K

 2  2pvb from   80 A:   A E T K A F L A D G D K D G D G M I G V D E F A A M I K A _ _

 3  2scp from    6 A:   Q K M K T Y F N R I D F D K D G A I T R M D F E S M A E R F A

 4  2sas from    9   :  K I K F T F D F F L D M N H D G S I Q D N D F E D M M T R Y K

 5  1cdl from   46 B:   A E L Q D M I N E V D A D G N G T I D F P E F L T M M A R K M

 6  1cdl from   46 C:   A E L Q D M I N E V D A D G N G T I D F P E F L T M M A R K M

 7  1alv from  170 B:   K K W Q A I Y K Q F D V D R S G T I G S S E L P G A F E A A G

 8  1alv from  140 A:   D T C R S M V A V M D S D T T G K L G F E E F K Y L W N N I K

 9  2pvb from   41 A:   D D V K K A F Y V I D Q D K S G F I E E D E L K L F L Q N F S

10  2sas from  105   :  N R I P F L F K G M D V S G D G I V D L E E F Q N Y C K N F Q

11  2cbl from  219 A:   L E A M A L K S T I D L T C N D Y I S V F E F D I F T R L F Q


  aa types - acidic, etc **

   relseq             -10         -5          0          5         10         15

 1  1acc from  167   :  _ _ _ _ _ _ _ _ h h a b a p a g h h a s h a h a g h s h a h b

 2  2pvb from   80 A:   s a s b s h h s a g a b a g a g h h g h a a h s s h h b s _ _

 3  2scp from    6 A:   p b h b s h h p b h a h a b a g s h s b h a h a s h s a b h s

 4  2sas from    9   :  b h b h s h a h h h a h p p a g s h p a p a h a a h h s b h b

 5  1cdl from   46 B:   s a h p a h h p a h a s a g p g s h a h h a h h s h h s b b h

 6  1cdl from   46 C:   s a h p a h h p a h a s a g p g s h a h h a h h s h h s b b h

 7  1alv from  170 B:   b b h p s h h b p h a h a b s g s h g s s a h h g s h a s s g

 8  1alv from  140 A:   a s h b s h h s h h a s a s s g b h g h a a h b h h h p p h b

 9  2pvb from   41 A:   a a h b b s h h h h a p a b s g h h a a a a h b h h h p p h s

10  2sas from  105   :  p b h h h h h b g h a h s g a g h h a h a a h p p h h b p h p
```

```
11  2cbl from  219 A:  h a s h s h b s s h a h s h p a h h s h h a h a h h s b h h p


   conformation sequences **

     relseq               -10          -5           0           5          10          15

 1  1acc from  167  :  . . . . . . . . . b d a k g k g b b a a a a a a . b b b b b b

 2  2pvb from   80 A:  a a a a a a a a a a d k k g k g b b a a a a a a a a k . . .

 3  2scp from    6 A:  a a a a a a a a a k d a k g k g b b b a a a a a a a a a a a a

 4  2sas from    9  :  a a a a a a a a k . d a k g k g b b b a a a a a a a a a a a a

 5  1cdl from   46 B:  a a a a k a a a a a b a k g a g b b b a a a a a a a a k b b b

 6  1cdl from   46 C:  a a a a a a a a a a b a k g k g b b b a a a a a a a a k b b b

 7  1alv from  170 B:  a a a a a a a a a k d k k g k g b b b a a a a a a a a a a k g

 8  1alv from  140 A:  a a a a a a a a a k d a k g k g b b b a a a a a a a a a a a a

 9  2pvb from   41 A:  a a a a a a a a a a d a k g k g b b b a a a a a k a a a k a b

10  2sas from  105  :  a a a a a a a k k a d a k g k g b b b a a a a a a a a a d . d

11  2cbl from  219 A:  a a a a a a a a a a d a k g k g b b b a a a a a a a a a a d b


 **********  protein names (pdb header) *******

 1  1acc at  177      2.1       TOXIN                Anthrax protective antigen

 2  2pvb at   90 A    0.91      METAL BINDING PROTEIN     EF Hand-like    Parvalbumin

 3  2scp at   16 A    2.0       BINDING PROTEIN             EF Hand-like   Calmodulin-like

 4  2sas at   19      2.4       CALCIUM-BINDING PROTEIN    EF Hand-like Calmodulin-like

 5  1cdl at   56 B    2.2       CALCIUM-BINDING PROTEIN    EF Hand-like   Calmodulin-like

 6  1cdl at   56 C    2.2       CALCIUM-BINDING PROTEIN    EF Hand-like   Calmodulin-like

 7  1alv at  180 B    1.9       CALCIUM BINDING      EF Hand-like EF-hand modules in multidomain proteins

 8  1alv at  150 A    1.9       CALCIUM BINDING

 9  2pvb at   51 A    0.91      METAL BINDING PROTEIN    EF Hand-like

10  2sas at  115      2.4       CALCIUM-BINDING PROTEIN

11  2cbl at  229 A    2.1       COMPLEX (PROTO-ONCOGENE/PEPTIDE)    EF-hand modules in multidomain proteins


 **********  cngps  ************

 1 1   DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1acc  1 0.2 2.1  CA  800 ASP  177 .....    |     Z

 2 1   DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 2pvb  0 0.1 0.9  CA  111 ASPA  90 ....b    |     Z

 3 1   DDDOD     Ca   2   2   2   5  -1  -1  -1 ..... 6 2scp  0 0.1 2.0  CA  190 ASPA  16 ....b    |     Z
```

```
  4  1   DNDOD      Ca    2    2    2    5   -1   -1   -1 .....  5  2sas   0  0.2  2.4   CA   186  ASP    19  ....b    |
  5  1   DDNNOE    Ca    2    2    0    2    5   -1   -1 ......6  1cdl   0  0.6  2.2   CA     2  ASPB   56  .....b   |
  6  1   DDNODE    Ca    2    2    2    2    3   -1   -1 ......6  1cdl   0  0.6  2.2   CA     2  ASPC   56  .....b   |
  7  1   DDSOE      Ca    2    2    2    5   -1   -1   -1 .....  6  1alv   0  0.1  1.9   CA     7  ASPB  180  ....b    |     Z   3.
  8  1   DDTOE      Ca    2    2    2    5   -1   -1   -1 .....  6  1alv   0  0.2  1.9   CA     2  ASPA  150  ....b    |     Z   3.
  9  1   DDSOEE    Ca    2    2    2    2    3   -1   -1 ......6  2pvb   0  0.1  0.9   CA   110  ASPA   51  .....b   |
 10  1   DSDOE      Ca    2    2    2    5   -1   -1   -1 .....  5  2sas   0  0.2  2.4   CA   188  ASP   115  ....b    |
 11  1   DTNOE      Ca    2    2    2    5   -1   -1   -1 .....  6  2cbl   0  0.1  2.1   CA   352  ASPA  229  ....b    |     Z
```

p chain 1, 1acc, is different

all others:

over fi 1-11, psi 0-10 all agree with model (no 2) within 22 deg

```
 relseq      meanfi           meanpsi       number

      0         -76( 10)          75(  7)          10         d
      1         -61(  9)         -33( 14)          10         a
      2         -89( 12)           3(  7)          10         k
      3          59( 13)          32( 18)          10         g
      4         -91( 18)           0( 13)          10         k
      5          82( 12)           3( 12)          10         g
      6        -136(  8)         152( 14)          10         b
      7         -95( 11)         121( 10)          10         b
      8        -101( 28)         171(  8)          10         b
      9         -56(  5)         -39(  8)          10         a
     10         -63(  8)         -35(  8)          10         a
     11         -68(  8)         -37(  8)          10         a
```

local conformations pretty similar over relseq -10 to17, but diverge beyond
   that except that even there 3,4,8 are alike.

rms agreement of fi and psi together

```
pchain        3   4   5  11   2   6   7   8   9  10
```

```
   3  2scp     0  10  15  22  21  14  21  19  17  19

   4  2sas    10   0  19  21  21  16  22  21  19  21

   5  1cdl    15  19   0  27  21  15  22  17  15  19

  11  2cbl    22  21  27   0  23  19  17  19  19  22

   2  2pvb    21  21  21  23   0  13  15  15  13  13

   6  1cdl    14  16  15  19  13   0  15  14  11  14

   7  1alv    21  22  22  17  15  15   0  11   9  17

   8  1alv    19  21  17  19  15  14  11   0   8  17

   9  2pvb    17  19  15  19  13  11   9   8   0  13

  10  2sas    19  21  19  22  13  14  17  17  13   0

pchain          3   4   5  11   2   6   7   8   9  10
```

```
DDOO 2 2 2

    amino-acid sequences ****
      start of selected part of cngroup        |
      relseq                -10         -5          0          5         10         15
  1  1fzc from  308 C:  N G M Q F S T W D N D N D K F E G N C A E Q D G S G
  2  3fib from  308  :  N G M Q F S T W D N D N D K F E G N C A E Q D G S G

    aa types - acidic, etc **
      relseq                -10         -5          0          5         10         15
  1  1fzc from  308 C:  p g h p h s s h a p a p a b h a g p h s a p a g s g
  2  3fib from  308  :  p g h p h s s h a p a p a b h a g p h s a p a g s g

    conformation sequences **
      relseq                -10         -5          0          5         10         15
  1  1fzc from  308 C:  b g b b b b b a k . b d b k b a . b a a a a k g b b
  2  3fib from  308  :  b g b b b b b a k b b k b k b a . b a a a a a g . b

  **********  protein names (pdb header) *******
  1  1fzc at  318 C   2.3      BLOOD COAGULATION   Fibrinogen C-terminal domain-like
  2  3fib at  318     2.1      BLOOD COAGULATION   Fibrinogen C-terminal domain-like

  **********  cngps  ************
  1 1  DDOO     Ca   2   2   2  -1  -1  -1  -1 ....  4 1fzc  0 0.2 2.3  CA    1 ASPC 318 b...    |
  2 1  DDOO     Ca   2   2   2  -1  -1  -1  -1 ....  6 3fib  0 0.2 2.1  CA  400 ASP  318 b...    |   ZZ

within cngp same with s.d. 1-9 deg
    relseq     meanfi          meanpsi       number
        0        -84(  1)        120(  4)         2        b
        1       -135(  2)         28(  6)         2        k
        2        -86(  2)        182(  9)         2        b
        3        -99(  8)         19(  3)         2        k
```

```
      4        -104(   2)         150(   5)            2         b
      5          -67(   4)         -33(   1)            2         a
      6          -90(   1)       -156(   0)            2         .
```

local conformations same (r.m.s. 8 deg)

proteins same fold and family

first D is bidentate

coordination no 4 or 6 - latter, in 3fib, has 2 waters

```
Zn CCCC 3 3 8       (in 90% cull set of proteins, protein chains 4 and 5 are in 30% set)


**********  cngps  ************

 1 1  CCCC      Zn   3   3   8  -1  -1  -1  -1 ....  4 1cdo  0 0.2 2.0  ZN  377 CYSA  98 ....    |        1.

 2 1  CCCC      Zn   3   3   8  -1  -1  -1  -1 ....  4 1d1t  0 0.2 2.4  ZN  375 CYSA  97 ....    |        1.

 3 1  CCCC      Zn   3   3   8  -1  -1  -1  -1 ....  4 1e3i  0 0.1 2.1  ZN  381 CYSA  97 ....    |        1.

 4 1  CCCC      Zn   3   3   8  -1  -1  -1  -1 ....  4 1e3j  0 0.2 2.3  ZN  901 CYSA  96 ....    |

 5 1  CCCC      Zn   3   3   8  -1  -1  -1  -1 ....  4 1het  0 0.2 1.1  ZN  401 CYSA  97 ....    |        1.

 6 1  CCCC      Zn   3   3   8  -1  -1  -1  -1 ....  4 1ht0  0 0.2 2.0  ZN 1375 CYSA  97 ....    |        1.


  amino-acid sequences ****

   relseq              -10         -5          0          5         10         15

 1  1cdo from   88 A:  E K V I P L F I S Q C G E C R F C Q S P K T N Q C V K G W A N E S P D V M S P K E T R F T C

 2  1d1t from   87 A:  D K V I P L F L P Q C R E C N A C R N P D G N L C I R S D _ _ _ G R G V L A D G T T R F T C

 3  1e3i from   87 A:  D K V I P F F A P Q C K R C K L C L S P L T N L C G K L R N F K Y P T I D Q E L M E D R T S

 4  1e3j from   86 A:  D R V A V E P G V P C R R C Q F C K E G K Y N L C P D L T F C A T P P D D G N L A R Y Y V H

 5  1het from   87 A:  D K V I P L F T P Q C G K C R V C K H P E G N F C L K N D L S M P R G T M Q D G T S R F T C

 6  1ht0 from   87 A:  D K V I P L F T P Q C G K C R I C K N P E S N Y C L K N D L G N P R G T L Q D G T R R F T C


  conformation sequences  (Hovmoller type) **

   relseq              -10         -5          0          5         10         15

 1  1cdo from   88 A:  b b b b b b a b b b a g a b a a a a b a a b a b b a b j a a a a b a a a g b b a b b a b b b

 2  1d1t from   87 A:  b b b b b b a b b b a g a b a a a a b a a b a b b a a b b b b b g a g b b a a g b b a b b b

 3  1e3i from   87 A:  b b b b b b a b b b a g a b a a a a b a a b a b b a a b a b a a b a a a b g a b b a a g b b

 4  1e3j from   86 A:  b b b b b b a b b b a g a b a a a a a g a a a a b a a b b b b g b b a b b b a b b a b b b b

 5  1het from   87 A:  b b b b b b a b b b a g a b a a a a b a a b a b b a a b a a a b b a g b b a a g b b a b b b

 6  1ht0 from   87 A:  b b b b b b a b b b a g a b a a a a b a a b a b b a a b a a a b b a g b b a a g b b a b b b


**********  protein names (pdb header) *******

 1  1cdo at   98 A   2.05     OXIDOREDUCTASE (CH-OH(D)-NAD(A))

 2  1d1t at   97 A   2.40     OXIDOREDUCTASE

 3  1e3i at   97 A   2.08     ALCOHOL DEHYDROGENASE
```

```
  4  1e3j at   96 A   2.3      OXIDOREDUCTASE

  5  1het at   97 A   1.15     OXIDOREDUCTASE

  6  1ht0 at   97 A   2.00     OXIDOREDUCTASE


  **********  SCOP classification *******

  1  1cdo at   98 A   2.05    GroES-like /Alcohol dehydrogenase-like, N-terminal domain

  2  1d1t at   97 A   2.40    same

  3  1e3i at   97 A   2.08    same

  4  1e3j at   96 A   2.3      same

  5  1het at   97 A   1.15    same

  6  1ht0 at   97 A   2.00    same


over 1-14/0-13, i.e. whole group, 2,5,6 agree within 15 deg

                                          1,3, differ from them by 20-37 deg

                                          4 differs by 60 deg


      1-6/0-5,  i.e. first two loops, 1,2,4,5,6 agree within 12 deg.

                                                 3 is different by 23-30 deg


      7-14/6-13, i.e. 3rd loop, 2,5,6 agree within 18 deg

                                        1,3 agree with each other within 18 deg,

                                                (but 30-40 from 2,5,6)

                                        4 different by 70 deg from rest
```

```
Zn   CCCC    3   17    3
        all  except  protein  chain  no.  19  are  in  the  30%  cull  set  of  proteins;
                        protein  chain  19  is  only  in  the  90%  set.



**********   cngps   ************
17  1   CCCC       Zn    3   17    3   -1   -1   -1   -1  ....   4  1dcq   0  0.1  2.1   ZN   600  CYSA  264  ....     |
18  1   CCCC       Zn    3   17    3   -1   -1   -1   -1  ....   4  1ee8   0  0.1  1.9   ZN   501  CYSA  238  ....     |
19  1   CCCC       Zn    3   17    3   -1   -1   -1   -1  ....   4  1ffy   0  0.1  2.2   ZN  1001  CYSA  886  ....     |         6.
20  1   CCCC       Zn    3   17    3   -1   -1   -1   -1  ....   4  1rmd   0  0.1  2.1   ZN   118  CYS    26  ....     |
21  1   CCCC       Zn    3   17    3   -1   -1   -1   -1  ....   4  1zin   0  0.1  1.6   ZN   219  CYS   130  ....     |         2.

    amino-acid  sequences  ****
      relseq                -10          -5           0           5          10          15
17   1dcq  from   254  A:  E  V  Q  R  M  T  G  N  D  V  C  C  D  C  G  A  P  D  P  T  W  L  S  T  N  L  G  I  L  T  C  I  E  C  S  G  I  H  R  E  L  G  V  H  Y  S
18   1ee8  from   228  A:  H  A  V  Y  G  R  E  G  L  P  C  P  A  C  G  R  P  V  E  R  R  V  V  A  G  R  G  T  H  F  C  P  T  C  Q  G  E  G  _  _  _  _  _  _  _  _
19   1ffy  from   876  A:  I  V  I  E  H  A  D  G  E  K  C  E  R  C  W  N  Y  S  E  D  L  G  A  V  D  E  L  T  H  L  C  P  R  C  Q  Q  V  V  K  S  L  _  _  _  _  _
20   1rmd  from    16   :   F  P  A  H  F  V  K  S  I  S  C  Q  I  C  E  H  I  L  A  D  P  V  E  T  S  C  K  H  L  F  C  R  I  C  I  L  R  C  L  K  V  M  G  S  Y  C
21   1zin  from   120   :  L  M  E  R  L  T  G  R  R  I  C  R  N  C  G  A  T  Y  H  L  I  F  H  P  P  A  K  P  G  V  C  D  K  C  G  G  E  L  Y  Q  R  A  D  D  N  E

  conformation  sequences   (Hovmoller  type)  **
                             -10          -5           0           5          10          15          20          25
17   1dcq  from   254  A:  a  a  a  a  b  b  g  a  a  b  b  a  a  a  g  b  b  g  b  a  b  b  b  a  a  a  g  b  b  b  b  a  a  a  a  a  a  a  a  a  a  j  a  a  a  b
18   1ee8  from   228  A:  b  a  a  b  g  a  b  g  b  b  b  a  a  a  g  b  b  b  b  b  b  b  b  g  g  b  b  b  b  b  b  a  a  a  a  g  b  j  b  b  b  b  b  b  b  b
19   1ffy  from   876  A:  b  b  b  b  b  b  b  j  b  b  b  a  a  a  g  b  b  b  a  a  b  b  b  b  j  a  b  b  g  b  b  a  a  a  a  a  a  a  a  a  a  b  b  b  b  b
20   1rmd  from    16   :   b  b  a  a  a  a  a  a  a  b  b  a  a  a  g  b  b  b  b  g  b  b  b  b  a  a  g  b  b  b  b  a  a  a  a  a  a  a  a  a  a  a  j  b  b  b
21   1zin  from   120   :  a  a  a  a  a  a  a  b  b  b  b  a  a  a  g  b  b  b  b  a  a  a  b  b  b  a  b  b  g  b  b  a  a  a  g  b  b  b  b  b  b  a  a  a  b  a



**********   protein  names  (pdb  header)  *******
17   1dcq  at   264  A    2.10       METAL  BINDING  PROTEIN
18   1ee8  at   238  A    1.90       DNA  BINDING  PROTEIN
```

19 1ffy at 886 A 2.2 LIGASE/RNA

20 1rmd at 26 2.1 DNA-BINDING PROTEIN

21 1zin at 130 1.6 PHOSPHOTRANSFERASE

over -1 to 7, i.e. first loop +, 18,19,20,21 agree, within 16

17 a bit different

over -1 to 23, i.e. all three loops p17-p21 are ALL DIFFERENT - by 40-90 deg !

NOTE THAT THE SEQUENCES SHOW MUCH LESS SIMILARITY FOR THIS SET, AND THE FOLDS ARE ALL DIF.

********** SCOP classification *******

17 1dcq at 264 A 2.10 Pyk2-associated protein beta ARF-GAP domain

18 1ee8 at 238 A 1.90 DNA repair protein MutM (Fpg

19 1ffy at 886 A 2.2 Anticodon-binding domain of a subclass of class I aminoacyl-tRNA synthetases

20 1rmd at 26 2.1 RING finger domain, C3HC4

21 1zin at 130 1.6 Rubredoxin-like /Microbial and mitochondrial ADK, insert "zinc finger" domain

recheck fold info from scratch (i.e. start with findseq) 24.4.02

1dcq Pyk2-associated protein beta ARF-GAP domain

Class: Small proteins

1ee8 DNA repair protein MutM (Fpg)

Class: Multi-domain proteins (alpha and beta)

1ffy Anticodon-binding domain of a subclass of class I aminoacyl-tRNA synthetases

Class: All alpha proteins

1rmd RING finger domain, C3HC4

Class: Small proteins

1zin Rubredoxin-like Microbial and mitochondrial ADK, insert "zinc finger" domain

Class: Small proteins

```
repeat torag to get agreements


loop 1

======                5 protein chains to be compared

start and end seq values for fi    1    3; for psi     0    2


   rms agreement of fi and psi together

pchain       17 18 19 20 21


 17 1dcq      0 20 29 35 23


 18 1ee8    20   0 19 23 16


 19 1ffy    29 19   0   6 11


 20 1rmd    35 23   6   0 16


 21 1zin    23 16 11 16   0



loop 2

======

                5 protein chains to be compared

start and end seq values for fi    4   20; for psi     3   19


   rms agreement of fi and psi together

pchain       17 18 19 20 21


 17 1dcq      0 64 75 52 90


 18 1ee8    64   0 69 63 84


 19 1ffy    75 69   0 67 63
```

```
  20 1rmd    52 63 67  0 82


  21 1zin    90 84 63 82  0



loop 3
======
                5 protein chains to be compared
start and end seq values for fi  21  23; for psi   20  22


   rms agreement of fi and psi together
pchain       17 18 19 20 21


  17 1dcq     0 24  7 19 32


  18 1ee8    24  0 21 16 27


  19 1ffy     7 21  0 16 28


  20 1rmd    19 16 16  0 27


  21 1zin    32 27 28 27  0
```

```
Zn HHH 4 6
 (in 90% cull set of proteins, protein chains 7 and 12  are in 30% set)
 ( redundant chains have not been removed)

includes some related cngroups  HHHXX 4 6 + + where additional residues
  in a second protein chain are coordinated to Zn

 18 groups and ALL agree over relseq -6 to 10 within <= 13 deg.

  amino-acid sequences ****
   start of selected part of cngroup      |
                        -10        -5          0          5         10         15
 1  1bud from 132 A:  V N R L V A I T L A H E M A H N L G V S H D E G S C S C G G
 2  1iag from 132  :  I N L L V A V T M A H E L G H N L G M E H D G K D C L R G A
 3  1qua from 132 A:  I P L L M A V T M A H E L G H N L G M N H D G A G C S C A T
 4  1sat from 166  :  T E D Y G R Q T F T H E I G H A L G L S H P G D Y N A G E G
 5  1ciz from 191 A:  T G T N L F L V A A H E I G H S L G L F H S A N T E A L M Y
 6  1hfc from 208  :  R E Y N L H R V A A H E L G H S L G L S H S T D I G A L M Y
 7  1i76 from 187 A:  A N Y N L F L V A A H E F G H S L G L A H S S D P G A L M Y
 8  1mmq from 208  :  L G I N F L Y A A T H E L G H S L G M G H S S D P N A V M Y
 9  830c from 212 A:  K G Y N L F L V A A H E F G H S L G L D H S K D P G A L M F
10  830c from 212 B:  K G Y N L F L V A A H E F G H S L G L D H S K D P G A L M F
11  1fbl from 208  :  R D Y N L Y R V A A H E L G H S L G L S H S T D I G A L M Y
12  1bkc from 395 E:  L T K E A D L V T T H E L G H N F G A E H D P D G K A E C A
13  1bkc from 395 A:  L T K E A D L V T T H E L G H N F G A E H D P D G L A E C A
14  1bkc from 395 I:  L T K E A D L V T T H E L G H N F G A E H D P D G L A E C A
15  1bkc from 395 C:  L T K E A D L V T T H E L G H N F G A E H D P D G L A E C A
16  1kap from 166 P:  N G N Y G R Q T L T H E I G H T L G L S H P G D Y N A G E G
17  1atl from 132 A:  I N L L M G V T M A H E L G H N L G M E H D G K D C L R G A
18  1atl from 132 B:  I N L L M G V T M A H E L G H N L G M E H D G K D C L R G A

  conformation sequences  (Hovmoller type) **
```

```
   start of selected part of cngroup      |
                         -10         -5         0          5          10         15
 1  1bud from  132 A:  b a a a a a a a a a a a a a a a a g b b b b b j a b b a g j
 2  1iag from  132  :  b a a a a a a a a a a a a a a a a g b b b b b a a b b b j a
 3  1qua from  132 A:  b a a a a a a a a a a a a a a a a g b b b b b a a a g b a a
 4  1sat from  166  :  a b a a a a a a a a a a a a a a a g b b b a b b b b b g a b
 5  1ciz from  191 A:  a b b b a a a a a a a a a a a a a g b b b b a b a a b a a b
 6  1hfc from  208  :  b b b b a a a a a a a a a a a a a g b b b b a b a a b a a b
 7  1i76 from  187 A:  b b b b a a a a a a a a a a a a a g b b b b a b a a b a a b
 8  1mmq from  208  :  b j b b a a a a a a a a a a a a a g b b b b a b a a b a a b
 9  830c from  212 A:  b j b b a a a a a a a a a a a a a g b b b b a b a a b a a b
10  830c from  212 B:  b j b b a a a a a a a a a a a a a g b b b b a b a a b a a b
11  1fbl from  208  :  b b g b a a a a a a a a a a a a a g b b b b a b a a b a a b
12  1bkc from  395 E:  b a a a a a a a a a a a a a a a a g b b b b b b g g a a a a
13  1bkc from  395 A:  b a a a a a a a a a a a a a a a a g b b b b b b j a a g a b
14  1bkc from  395 I:  b a a a a a a a a a a a a a a a a g b b b b b b j a b a a b
15  1bkc from  395 C:  b a a a a a a a a a a a a a a a a g b b b b b b j b j a a b
16  1kap from  166 P:  b g b a a a a a a a a a a a a a a g b b b a b b b b b g b j
17  1atl from  132 A:  b a a a a a a a a a a a a a a a a g b b b b b a a b b b j a
18  1atl from  132 B:  b a a a a a a a a a a a a a a a a g b b b b b a a b b b j a


**********  protein names (pdb header) + CATH and SCOP classification *******
 1  1bud at  142 A   1.90    TOXIN                      d.92.1.9 Zincin-like
 2  1iag at  142     2.0     METALLOPROTEASE         d.92.1.9
 3  1qua at  142 A   2.2     TOXIN
 4  1sat at  176     1.75    HYDROLASE (SERINE PROTEASE)
 5  1ciz at  201 A   1.64    METALLOPROTEINASE
 6  1hfc at  218     1.56    METALLOPROTEASE
 7  1i76 at  197 A   1.20    HYDROLASE               (no code given)
 8  1mmq at  218     1.9     METALLOPROTEASE
 9  830c at  222 A   1.6     MATRIX METALLOPROTEASE   d.92.1.11
10  830c at  222 B   1.6     MATRIX METALLOPROTEASE
11  1fbl at  218     2.5     METALLOPROTEASE
```

```
12  1bkc at  405 E   2.0       ZN-ENDOPEPTIDASE  d.92.1.10 Zincin-like
13  1bkc at  405 A   2.0       ZN-ENDOPEPTIDASE
14  1bkc at  405 I   2.0       ZN-ENDOPEPTIDASE
15  1bkc at  405 C   2.0       ZN-ENDOPEPTIDASE
16  1kap at  176 P   1.64     ZINC METALLOPROTEASE   d.92.1.6  Zincin-like
17  1atl at  142 A   1.8       METALLOENDOPEPTIDASE   d.92.1.9  Zincin-like
18  1atl at  142 B   1.8       METALLOENDOPEPTIDASE

**********  cngps  ************
 1 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   4 1bud  0 0.1 1.9  ZN  800 HISA 142 ...      |  U
 2 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   4 1iag  0 0.2 2.0  ZN  999 HIS  142 ...      |  Z     3.
 3 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   4 1qua  0 0.1 2.2  ZN  999 HISA 142 ...      |  Z
 4 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   4 1sat  1 0.2 1.8  ZN  472 HIS  176 ...      |  Z
 5 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   5 1ciz  0 0.3 1.6  ZN  301 HISA 201 ...      |  UU   3.
 6 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   5 1hfc  0 0.1 1.6  ZN  275 HIS  218 ...      |  UU   3.
 7 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   5 1i76  0 0.3 1.2  ZN  999 HISA 197 ...      |  UU   3.
 8 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   5 1mmq  1 0.2 1.9  ZN    1 HIS  218 ...      |  UU   3.
 9 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   5 830c  1 0.2 1.6  ZN  261 HISA 222 ...      |  UU   3.
10 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   5 830c  1 0.2 1.6  ZN  261 HISB 222 ...      |  UU   3.
11 1  HHH        Zn   4   6  -1  -1  -1  -1  -1 eee   6 1fbl  0 0.5 2.5  ZN  998 HIS  218 ...      |  BUU  3.
12 1  HHH O     Zn   4   6  -1-413  -1  -1  -1 eee . 5 1bkc  0 0.4 2.0  ZN    1 HISE 405 ... .   |  U
13 1  HHH O     Zn   4   6  -1-413  -1  -1  -1 eee . 5 1bkc  1 0.4 2.0  ZN    1 HISA 405 ... .   |  U
14 1  HHH O     Zn   4   6  -1-413  -1  -1  -1 eee . 5 1bkc  1 0.4 2.0  ZN    1 HISI 405 ... .   |  U
15 1  HHH O     Zn   4   6  -1-413  -1  -1  -1 eee . 5 1bkc  1 0.5 2.0  ZN    1 HISC 405 ... .   |  U
16 1  HHHOS     Zn   4   6 -99   0  -1  -1  -1 eee.. 5 1kap  0 0.2 1.6  ZN  613 HISP 176 .....   |        3.
17 1  HHHOY     Zn   4   6 -99   0  -1  -1  -1 eee.. 5 1atl  0 0.1 1.8  ZN  401 HISA 142 .....   |        3.
18 1  HHHOY     Zn   4   6 -99   0  -1  -1  -1 eee.. 5 1atl  0 0.1 1.8  ZN  402 HISB 142 .....   |        3.

 18 groups and ALL agree over relseq -6 to 10 within <= 13 deg.

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

Separate analysis of simple HHH groups in 90% cull set (different numbering)

```
 5  1bud from  132 A:  V N R L V A I T L A H E M A H N L G V S H D E G S C S C G G K S C I
 6  1iag from  132  :  I N L L V A V T M A H E L G H N L G M E H D G K D C L R G A S L C I
 7  1qua from  132 A:  I P L L M A V T M A H E L G H N L G M N H D G A G C S C A T C I M A
 8  1sat from  166  :  T E D Y G R Q T F T H E I G H A L G L S H P G D Y N A G E G D P T Y
 9  1ciz from  191 A:  T G T N L F L V A A H E I G H S L G L F H S A N T E A L M Y P L Y H
10  1hfc from  208  :  R E Y N L H R V A A H E L G H S L G L S H S T D I G A L M Y P S Y T
11  1i76 from  187 A:  A N Y N L F L V A A H E F G H S L G L A H S S D P G A L M Y P N Y A
12  1mmq from  208  :  L G I N F L Y A A T H E L G H S L G M G H S S D P N A V M Y P T Y G
13  830c from  212 A:  K G Y N L F L V A A H E F G H S L G L D H S K D P G A L M F P I Y T
14  830c from  212 B:  K G Y N L F L V A A H E F G H S L G L D H S K D P G A L M F P I Y T
15  1fbl from  208  :  R D Y N L Y R V A A H E L G H S L G L S H S T D I G A L M Y P N Y I


 5  1bud from  132 A:  b a a a a a a a a a a a a a a a a g b b b b b j a b b a g j a b b a
 6  1iag from  132  :  b a a a a a a a a a a a a a a a a g b b b b b a a b b b j a b b b a
 7  1qua from  132 A:  b a a a a a a a a a a a a a a a a g b b b b b a a a g b a a b a a b
 8  1sat from  166  :  a b a a a a a a a a a a a a a a a g b b b a b b b b b g a b b b b a
 9  1ciz from  191 A:  a b b b a a a a a a a a a a a a a g b b b b a b a a b a a b a b b b
10  1hfc from  208  :  b b b b a a a a a a a a a a a a a g b b b b a b a a b a a b a b b b
11  1i76 from  187 A:  b b b b a a a a a a a a a a a a a g b b b b a b a a b a a b a b b b
12  1mmq from  208  :  b j b b a a a a a a a a a a a a a g b b b b a b a a b a a b a b b g
13  830c from  212 A:  b j b b a a a a a a a a a a a a a g b b b b a b a a b a a b a b b b
14  830c from  212 B:  b j b b a a a a a a a a a a a a a g b b b b a b a a b a a b a b b b
15  1fbl from  208  :  b b g b a a a a a a a a a a a a a g b b b b a b a a b a a b a b b b
                        -10         -5          0          5         10         15


                ------------0-----4-----------10-------
                                H       H                H
```

10 and 11 same over whole range within 6 degs

13 and 14 same over whole range within 14 degs

all same over relseq -6 to 10 within 13 degs, many much closer

1i76 at 197 A should be model

```
 5  1bud  1.90        TOXIN
 6  1iag  2.0          METALLOPROTEASE
 7  1qua  2.2          TOXIN
 8  1sat  1.75        HYDROLASE (SERINE PROTEASE)
 9  1ciz  1.64        METALLOPROTEINASE
10  1hfc  1.56        METALLOPROTEASE
11  1i76  1.20        HYDROLASE
12  1mmq  1.9          METALLOPROTEASE
13  830c  1.6          MATRIX METALLOPROTEASE
14  830c  1.6          MATRIX METALLOPROTEASE
15  1fbl  2.5          METALLOPROTEASE

 5  1bud at  142 A:Snake venom metalloprotease from Five-pace snake       Zincin-like
 6  1iag at  142  :Snake venom metalloprotease from Eastern diamond       Zincin-like
 7  1qua at  142 A:  Snake venom metalloprotease from Chinese five        Zincin-like
 8  1sat at  176  : Metallo protease, catalytic (N-terminal) domain      Zincin-like
 9  1ciz at  201 A:Stromelysin-1 (MMP-3) from Human (Homo sapiens),      Zincin-like
10  1hfc at  218  :
11  1i76 at  197 A:
12  1mmq at  218  :
13  830c at  222 A:
14  830c at  222 B:
15  1fbl at  218  :


 5  1bud at  142 A:
 6  1iag at  142  :
 7  1qua at  142 A:
 8  1sat at  176  :
```

```
 9   1ciz  at   201  A:

10   1hfc  at   218   :

11   1i76  at   197  A:

12   1mmq  at   218   :

13   830c  at   222  A:

14   830c  at   222  B:

15   1fbl  at   218   :


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

COMMONLY OCCURRING CHELATE LOOPS in ZN and CA COORDINATION GROUPS

CHELATE LOOP: the term donor pair was used previously for this.

The data includes, for the individual occurrences:

i) amino-acid sequences, starting at *relseq*=-10; *relseq*=0 is the first donor amino-acid

ii) sequences of amino-acid types - g is glycine, s is small hydrophobic, h is large hydrophobic, a is acidic, b is basic, p is polar

iii) sequences of (Efimov type) conformations

iv) starting residue number, resolution, PDB header name,

v) coordination group definition (as in Table 3D)

vi) summary of analysis of agreement of torsion angles over selected regions of protein chain; chains are identified by their protein chain numbers (pc) - this is the first number in each row of the tables.

[Occasionally, additional information is given from an earlier analysis of donor pairs which included metal coordination geometry; the protein chain numbering may be different and the conformations are described by the categories of Hovmoller et al., using a for helix, b for sheet, g for turn, and j for other.]

Ca D D 2

Ca D N 2

Ca D O 2

Ca N O 2

Ca D O 1

Ca O D 0

Ca O D 2

Ca O E 2

Ca O E 5

Ca O S 0 and Ca O T 0

Ca O O 2

Ca O O 3

.

Zn H H 4

Zn C C 2

Zn C C 3

Zn C C 5

Zn H C 3

Zn H H 2

```
Ca D D 2


   amino-acid sequences ****
     start of selected part of cngroup          |
     relseq                 -10          -5           0           5          10          15
 1   2scp from     6 A:  Q K M K T Y F N R I D F D K D G A I T R M D
 2   1cdl from    10 A:  A E F K E A F S L F D K D G D G T I T T K E
 3   1g4y from    10 R:  A E F K E A F S L F D K D G D G T I T T K E
 4   1sra from   247  :  H C T T R F F E T C D L D N D K Y I A L D E
 5   1vrk from    10 A:  A E F K E A F S L F D K D G D G T I T T K E
 6   1cdl from    46 B:  A E L Q D M I N E V D A D G N G T I D F P E
 7   1rec from   100  :  Q K L E W A F S L Y D V D G N G T I S K N E
 8   1vrk from    46 A:  A E L Q D M I N E V D A D G N G T I D F P E
 9   2pvb from    41 A:  D D V K K A F Y V I D Q D K S G F I E E D E
10   1alv from   140 A:  D T C R S M V A V M D S D T T G K L G F E E
11   2pvb from    80 A:  A E T K A F L A D G D K D G D G M I G V D E
12   1g4y from    46 R:  A E L Q D M I N E V D A D G N G T I D F P E
13   1alv from   170 B:  K K W Q A I Y K Q F D V D R S G T I G S S E
14   1acc from   167  :  _ _ _ _ _ _ _ _ _ P D R D N D G I P D S L E
15   1e43 from   192 A:  N Y D Y L M Y A D V D Y D H P D V V A E T K
16   2scp from     8 A:  M K T Y F N R I D F D K D G A I T R M D F E
17   1cdl from    12 A:  F K E A F S L F D K D G D G T I T T K E L G
18   1g4y from    12 R:  F K E A F S L F D K D G D G T I T T K E L G
19   1sra from   249  :  T T R F F E T C D L D N D K Y I A L D E W A
20   1vrk from    12 A:  F K E A F S L F D K D G D G T I T T K E L G
21   2pvb from    82 A:  T K A F L A D G D K D G D G M I G V D E F A
22   1acc from   169  :  _ _ _ _ _ _ _ P D R D N D G I P D S L E V E
23   1acc from   169  :  _ _ _ _ _ _ _ P D R D N D G I P D S L E V E
24   1fzc from   308 C:  N G M Q F S T W D N D N D K F E G N C A E Q
25   3fib from   308  :  N G M Q F S T W D N D N D K F E G N C A E Q
26   1fzc from   371 B:  N G M F F S T Y D R D N D G W L T S D P R K
27   1alv from   213 B:  R R Y S D E G G N M D F D N F I S C L V R L
```

```
28  1alv from  213 A:  R R Y S D E G G N M D F D N F I S C L V R L
29  2por from   83  :  F G D L Y E V G Y T D L D D R G G N D I P Y


   aa types - acidic, etc **
    relseq                -10         -5          0          5         10         15
 1  2scp from     6 A:  p b h b s h h p b h a h a b a g s h s b h a
 2  1cdl from    10 A:  s a h b a s h s h h a b a g a g s h s s b a
 3  1g4y from    10 R:  s a h b a s h s h h a b a g a g s h s s b a
 4  1sra from   247  :  p h s s b h h a s h a h a p a b h h s h a a
 5  1vrk from    10 A:  s a h b a s h s h h a b a g a g s h s s b a
 6  1cdl from    46 B:  s a h p a h h p a h a s a g p g s h a h h a
 7  1rec from   100  :  p b h a h s h s h h a h a g p g s h s b p a
 8  1vrk from    46 A:  s a h p a h h p a h a s a g p g s h a h h a
 9  2pvb from    41 A:  a a h b b s h h h h a p a b s g h h a a a a
10  1alv from  140 A:  a s h b s h h s h h a s a s s g b h g h a a
11  2pvb from   80 A:  s a s b s h h s a g a b a g a g h h g h a a
12  1g4y from   46 R:  s a h p a h h p a h a s a g p g s h a h h a
13  1alv from  170 B:  b b h p s h h b p h a h a b s g s h g s s a
14  1acc from  167  :  _ _ _ _ _ _ _ _ _ h a b a p a g h h a s h a
15  1e43 from  192 A:  p h a h h h h s a h a h a p h a h h s a s b
16  2scp from    8 A:  h b s h h p b h a h a b a g s h s b h a h a
17  1cdl from   12 A:  h b a s h s h h a b a g a g s h s s b a h g
18  1g4y from   12 R:  h b a s h s h h a b a g a g s h s s b a h g
19  1sra from  249  :  s s b h h a s h a h a p a b h h s h a a h s
20  1vrk from   12 A:  h b a s h s h h a b a g a g s h s s b a h g
21  2pvb from   82 A:  s b s h h s a g a b a g a g h h g h a a h s
22  1acc from  169  :  _ _ _ _ _ _ _ h a b a p a g h h a s h a h a
23  1acc from  169  :  _ _ _ _ _ _ _ h a b a p a g h h a s h a h a
24  1fzc from  308 C:  p g h p h s s h a p a p a b h a g p h s a p
25  3fib from  308  :  p g h p h s s h a p a p a b h a g p h s a p
26  1fzc from  371 B:  p g h h h s s h a b a p a g h h s s a h b b
27  1alv from  213 B:  b b h s a a g g p h a h a p h h s h h h b h
28  1alv from  213 A:  b b h s a a g g p h a h a p h h s h h h b h
```

```
29   2por from    83   :  h g a h h a h g h s a h a a b g g p a h h h


   conformation  sequences  **

     relseq                -10         -5          0          5          10          15

 1   2scp from     6  A:  a a a a a a a a a a b a a g a g b b b a a a

 2   1cdl from    10  A:  a a a a a a a a a a b a a g a g b b b a a a

 3   1g4y from    10  R:  a a a a a a a a a a b a a g a g b b b a a a

 4   1sra from   247   :  a a a a a a a a a a b a a g a g b b b a a a

 5   1vrk from    10  A:  a a a a a a a a a a b a a g a g b b b a a a

 6   1cdl from    46  B:  a a a a a a a a a a b a a g a g b b b a a a

 7   1rec from   100   :  a a a a a a a a a a b a a g a g b b b a a a

 8   1vrk from    46  A:  a a a a a a a a a a b a a g a g b b b a a a

 9   2pvb from    41  A:  a a a a a a a a a a b a a g a g b b b a a a

10   1alv from   140  A:  a a a a a a a a a a b a a g a g b b b a a a

11   2pvb from    80  A:  a a a a a a a a a a b a a g a g b b b a a a

12   1g4y from    46  R:  a a a a a a a a a a b a a g a g b b b a a a

13   1alv from   170  B:  a a a a a a a a a a b a a g a g b b b a a a

14   1acc from   167   :  b b b b b b b b b b b a a g a g b b a a a a

15   1e43 from   192  A:  b a a b a b j b b b b a a b a a a a a a a a

16   2scp from     8  A:  a a a a a a a a b a a g a g b b b a a a a a

17   1cdl from    12  A:  a a a a a a a a b a a g a g b b b a a a a a

18   1g4y from    12  R:  a a a a a a a a b a a g a g b b b a a a a a

19   1sra from   249   :  a a a a a a a a b a a g a g b b b a a a a a

20   1vrk from    12  A:  a a a a a a a a b a a g a g b b b a a a a a

21   2pvb from    82  A:  a a a a a a a a b a a g a g b b b a a a a a

22   1acc from   169   :  b b b b b b b b b a a g a g b b a a a a a a

23   1acc from   169   :  b b b b b b b b b a a g a g b b a a a a a a

24   1fzc from   308  C:  b g b b b b b a a b b a b a b a b b a a a a

25   3fib from   308   :  b g b b b b b a a b b a b a b a b b a a a a

26   1fzc from   371  B:  b g b b b b b a a b b a b a b b b a b a a b

27   1alv from   213  B:  a a a b b a a g b b b a a a a a a a a a a a

28   1alv from   213  A:  a a a a g a a g b b b a a a a a a a a a a a

29   2por from    83   :  a j b b b b b j a a g b a a a g j a b b b b
```

```
  conformation  sequences  **   revised  -  efimov

    relseq                -10          -5           0           5          10

      15

 1   2scp  from     6  A:   a  a  a  a  a  a  a  a  a  k  d  a  k  g  k  g  b  b  b  a  a  a

 2   1cdl  from    10  A:   a  a  a  a  a  a  a  a  a  a  d  a  k  g  k  g  b  b  b  a  a  a

 3   1g4y  from    10  R:   a  a  a  a  a  a  a  a  k  a  b  a  k  g  k  g  b  b  b  a  a  a

 4   1sra  from   247   :   k  k  a  a  a  a  a  a  a  k  d  a  k  g  k  g  b  b  b  a  a  a

 5   1vrk  from    10  A:   a  a  a  a  a  a  a  a  a  a  d  a  k  g  k  g  b  b  b  a  a  a

 6   1cdl  from    46  B:   a  a  a  a  k  a  a  a  a  a  b  a  k  g  a  g  b  b  b  a  a  a

 7   1rec  from   100   :   k  k  a  a  a  a  a  a  a  a  d  a  k  g  k  g  b  b  b  a  a  a

 8   1vrk  from    46  A:   a  a  a  a  a  a  a  a  a  k  d  a  k  g  k  g  b  b  b  a  a  a

 9   2pvb  from    41  A:   a  a  a  a  a  a  a  a  a  a  d  a  k  g  k  g  b  b  b  a  a  a

10   1alv  from   140  A:   a  a  a  a  a  a  a  a  a  k  d  a  k  g  k  g  b  b  b  a  a  a

11   2pvb  from    80  A:   a  a  a  a  a  a  a  a  a  a  d  k  k  g  k  g  b  b  b  a  a  a

12   1g4y  from    46  R:   a  a  a  a  a  a  a  a  a  d  b  a  k  g  k  g  b  b  b  a  a  a

13   1alv  from   170  B:   a  a  a  a  a  a  a  a  a  k  d  k  k  g  k  g  b  b  b  a  a  a

14   1acc  from   167   :   .  .  .  .  .  .  .  .  .  b  d  a  k  g  k  g  b  b  a  a  a  a

15   1e43  from   192  A:   b  k  a  b  a  b  .  b  b  b  b  k  a  b  a  a  a  a  a  a  a  a

16   2scp  from     8  A:   a  a  a  a  a  a  a  k  d  a  k  g  k  g  b  b  b  a  a  a  a  a

17   1cdl  from    12  A:   a  a  a  a  a  a  a  a  d  a  k  g  k  g  b  b  b  a  a  a  a  a

18   1g4y  from    12  R:   a  a  a  a  a  a  k  a  b  a  k  g  k  g  b  b  b  a  a  a  a  a

19   1sra  from   249   :   a  a  a  a  a  a  a  k  d  a  k  g  k  g  b  b  b  a  a  a  a  a

20   1vrk  from    12  A:   a  a  a  a  a  a  a  a  d  a  k  g  k  g  b  b  b  a  a  a  a  a

21   2pvb  from    82  A:   a  a  a  a  a  a  a  a  d  k  k  g  k  g  b  b  b  a  a  a  a  a

22   1acc  from   169   :   .  .  .  .  .  .  .  b  d  a  k  g  k  g  b  b  a  a  a  a  a  a

23   1acc  from   169   :   .  .  .  .  .  .  .  b  d  a  k  g  k  g  b  b  a  a  a  a  a  a

24   1fzc  from   308  C:   b  g  b  b  b  b  b  a  k  .  b  d  b  k  b  a  .  b  a  a  a  a

25   3fib  from   308   :   b  g  b  b  b  b  b  a  k  b  b  k  b  k  b  a  .  b  a  a  a  a

26   1fzc  from   371  B:   b  g  b  b  b  b  b  k  k  b  b  k  b  k  b  b  b  a  b  a  a  d

27   1alv  from   213  B:   a  a  a  b  b  a  k  g  b  b  b  a  a  a  a  a  a  a  a  a  a  a

28   1alv  from   213  A:   a  a  k  a  g  .  k  g  b  b  b  a  a  a  a  a  a  a  a  a  a  a

29   2por  from    83   :   k  j  b  b  b  b  b  j  a  k  g  d  k  a  k  .  j  k  b  b  d  b
```

```
 **********  protein names (pdb header) *******

 1  2scp at   16 A   2.0       BINDING PROTEIN

 2  1cdl at   20 A   2.2       CALCIUM-BINDING PROTEIN

 3  1g4y at   20 R   1.6       SIGNALING PROTEIN

 4  1sra at  257     2.0       CALCIUM-BINDING PROTEIN

 5  1vrk at   20 A   1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT

 6  1cdl at   56 B   2.2       CALCIUM-BINDING PROTEIN

 7  1rec at  110     1.9       CALCIUM-BINDING PROTEIN

 8  1vrk at   56 A   1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT

 9  2pvb at   51 A   0.91      METAL BINDING PROTEIN

10  1alv at  150 A   1.9       CALCIUM BINDING

11  2pvb at   90 A   0.91      METAL BINDING PROTEIN

12  1g4y at   56 R   1.6       SIGNALING PROTEIN

13  1alv at  180 B   1.9       CALCIUM BINDING

14  1acc at  177     2.1       TOXIN

15  1e43 at  202 A   1.7       HYDROLASE

16  2scp at   18 A   2.0       BINDING PROTEIN

17  1cdl at   22 A   2.2       CALCIUM-BINDING PROTEIN

18  1g4y at   22 R   1.6       SIGNALING PROTEIN

19  1sra at  259     2.0       CALCIUM-BINDING PROTEIN

20  1vrk at   22 A   1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT

21  2pvb at   92 A   0.91      METAL BINDING PROTEIN

22  1acc at  179     2.1       TOXIN

23  1acc at  179     2.1       TOXIN

24  1fzc at  318 C   2.3       BLOOD COAGULATION

25  3fib at  318     2.1       BLOOD COAGULATION

26  1fzc at  381 B   2.3       BLOOD COAGULATION

27  1alv at  223 B   1.9       CALCIUM BINDING

28  1alv at  223 A   1.9       CALCIUM BINDING

29  2por at   93     1.8       INTEGRAL MEMBRANE PROTEIN PORIN
```

```
 **********  cngps  ************

 1 1  DDDOD     Ca   2   2   2   5  -1  -1  -1 ..... 6 2scp  0 0.1 2.0  CA  190 ASPA  16 ....b   |     Z

 2 1  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  0 0.4 2.2  CA    1 ASPA  20 .b..b   |

 3 1  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1g4y  0 0.2 1.6  CA 1001 ASPR  20 ....b   |     Z

 4 1  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1sra  0 0.2 2.0  CA  302 ASP  257 ....b   |     Z

 5 1  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.1 1.9  CA  151 ASPA  20 ....b   |     Z  2.

 6 1  DDNNOE   Ca   2   2   0   2   5  -1  -1 ......6 1cdl  0 0.6 2.2  CA    2 ASPB  56 .....b  |

 7 1  DDNOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1rec  0 0.2 1.9  CA  501 ASP  110 ....b   |     Z

 8 1  DDNOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.1 1.9  CA  152 ASPA  56 ....b   |     Z  2.

 9 1  DDSOEE   Ca   2   2   2   2   3  -1  -1 ......6 2pvb  0 0.1 0.9  CA  110 ASPA  51 .....b  |

10 1  DDTOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1alv  0 0.2 1.9  CA    2 ASPA 150 ....b   |     Z  3.

11 1  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 2pvb  0 0.1 0.9  CA  111 ASPA  90 ....b   |     Z

12 1  DDNOE     Ca   2   2   2   5  -1  -1  -1 ..... 5 1g4y  0 0.4 1.6  CA 1002 ASPR  56 ....b   |

13 1  DDSOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1alv  0 0.1 1.9  CA    7 ASPB 180 ....b   |     Z  3.

14 1  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1acc  1 0.2 2.1  CA  800 ASP  177 .....   |     Z

15 4  DODDD     Ca  22   2  19   2  -1  -1  -1 ..... 6 1e43  0 0.1 1.7  CA  502 ASPA 159 b....   |     Z  3.

16 2  DDDOD     Ca   2   2   2   5  -1  -1  -1 ..... 6 2scp  0 0.1 2.0  CA  190 ASPA  16 ....b   |     Z

17 2  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  0 0.4 2.2  CA    1 ASPA  20 .b..b   |

18 2  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1g4y  0 0.2 1.6  CA 1001 ASPR  20 ....b   |     Z

19 2  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1sra  0 0.2 2.0  CA  302 ASP  257 ....b   |     Z

20 2  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.1 1.9  CA  151 ASPA  20 ....b   |     Z  2.

21 2  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 2pvb  0 0.1 0.9  CA  111 ASPA  90 ....b   |     Z

22 2  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 1acc  1 0.2 2.1  CA  800 ASP  177 .....   |     Z

23 1  DDEOOD   Ca   2   7  34   3  10  -1  -1 ......6 1acc  0 0.3 2.1  CA  801 ASP  179 ..b...  |

24 1  DDOO      Ca   2   2   2  -1  -1  -1  -1 ....  4 1fzc  0 0.2 2.3  CA    1 ASPC 318 b...    |

25 1  DDOO      Ca   2   2   2  -1  -1  -1  -1 ....  6 3fib  0 0.2 2.1  CA  400 ASP  318 b...    |   ZZ

26 1  DDO       Ca   2   2  -1  -1  -1  -1  -1 ...   5 1fzc  0 0.2 2.3  CA    2 ASPB 381 b..     |  ZZ

27 2  DDDN      Ca  88   2   1  -1  -1  -1  -1 ....  6 1alv  0 0.1 1.9  CA    8 ASPB 135 .bb.    |   ZZ  3.

28 2  DDDN      Ca  88   2   1  -1  -1  -1  -1 ....  6 1alv  0 0.2 1.9  CA    4 ASPA 135 .bb.    |   ZZ  3.

29 1  DDND      Ca   2   5   1  -1  -1  -1  -1 ....  6 2por  0 0.2 1.8  CA  303 ASP   93 bb..    |   ZZ


  27, 28 seem to be same group, 28 should have been eliminated
```

over loop: 0-2

group 1: 1-10, same within 20 deg over loop
+11-13 within 20-30 deg
+ 14 15 within 25-35 deg
model for all: 2pvb at 51 A 0.91 (9)
of these 1-13 are related proteins and similar over relseq -10 to 11
14 15 quite different
1-14 are first loop of a 2 2 5 sequence

group 2: 16-23, same within 16 deg
model : 2pvb at 92 A 0.91 (21)
of these 16-21 are related proteins and similar over relseq -10 to 11
22 23 quite different
16-22 are second loop of a 2 2 5 sequence

group 3: 24-26, same within 8 deg - closely related proteins
model: 3fib at 318 2.1 (25)
of these 24 25 are pretty similar over relseq -10 to 11, 26 is dif !
(others: 27 28, same, 29 dif from all)

means pc 1-15 model: 2pvb at 51 A 0.91 (9)

relseq	meanfi	meanpsi	number		
0	-84(19)	79(9)	15	d	(or b)
1	-59(10)	-37(21)	15	a	(or k)
2	-88(12)	1(10)	15	k	(or a)

means pc 1-10

relseq	meanfi	meanpsi	number	(for st devns see below) q
0	-76	77	10	d

```
      1        -56        -43         10         a

      2        -85          1          10         k


    means  pc   16-23    model:  2pvb  at    92  A    0.91    (21)

    relseq    meanfi    meanpsi    number

      0        -89(  11)          0(   7)           8         k

      1          62(   9)         25(   7)           8         g

      2        -85(   5)          0(  11)           8         k


   means  pc  24-26      model:  3fib  at   318      2.1     (25)

      relseq    meanfi    meanpsi    number

      0        -84(   1)        118(   4)           3         b   **

      1       -134(   2)         27(   5)           3         k        (or  d)

      2        -82(   6)        183(   6)           3         b   **




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

p  chains  1-15

    relseq  0:   chi1  172  to  194,   chi2   -3  to  34    -  all  but  one

    relseq  2:   chi1  51  to  72,     chi2  -4  to  31     -  all  but  three


p  chains  16-23

    relseq  0:   al  have  chi1  51  to  65,   chi2  1  to  31

           2                    58  to  82,   chi2  -52  to  22


p  chains  24-26    chis  close


.............................................................................

                                                       SCOP  FOLD,  FAMILY

(in  old  order)

  2scp  at    16  A    2.0       BINDING  PROTEIN       EF  Hand-like   Calmodulin-like

  1cdl  at    20  A    2.2       CALCIUM-BINDING  PROTEIN

  1g4y  at    20  R    1.6       SIGNALING  PROTEIN           EF  Hand-like      Calmodulin-like
```

```
   1sra at  257      2.0       CALCIUM-BINDING PROTEIN      EF Hand-like/ Osteonectin

   1vrk at   20 A    1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT          Calmodulin/Calmodulin

   1cdl at   56 B    2.2       CALCIUM-BINDING PROTEIN     EF Hand-like / Calmodulin-like

   1rec at  110      1.9       CALCIUM-BINDING PROTEIN     EF Hand-like / Calmodulin-like

   1vrk at   56 A    1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT Calmodulin/Calmodulin

   2pvb at   51 A    0.91      METAL BINDING PROTEIN       EF Hand-like  Parvalbumin

   1alv at  150 A   1.9        CALCIUM BINDING               EF Hand-like/EF-hand modules in multidomain proteins

   2pvb at   90 A    0.91      METAL BINDING PROTEIN         EF Hand-like  Parvalbumin

   1g4y at   56 R    1.6       SIGNALING PROTEIN             EF Hand-like       Calmodulin-like

   1alv at  180 B   1.9        CALCIUM BINDING             EF Hand-like EF-hand modules in multidomain proteins

   1acc at  177      2.1       TOXIN                             Anthrax protective antigen

   1e43 at  202 A   1.7        HYDROLASE                       TIM beta/alpha-barrel/ alpha-Amylases, N-terminal domain

   2scp at   18 A    2.0       BINDING PROTEIN              EF Hand-like  Calmodulin-like

   1cdl at   22 A    2.2       CALCIUM-BINDING PROTEIN     EF Hand-like/ Calmodulin-like

   1g4y at   22 R    1.6       SIGNALING PROTEIN             EF Hand-like/ Osteonectin

   1sra at  259      2.0       CALCIUM-BINDING PROTEIN   EF Hand-like/ Osteonectin

   1vrk at   22 A    1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT          Calmodulin/Calmodulin

   2pvb at   92 A    0.91      METAL BINDING PROTEIN     EF Hand-like/ Parvalbumin

   1acc at  179      2.1       TOXIN                         Anthrax protective antigen

   1acc at  179      2.1       TOXIN                           Anthrax protective antigen


   1fzc at  318 C   2.3        BLOOD COAGULATION           Fibrinogen C-terminal domain-like

   3fib at  318      2.1       BLOOD COAGULATION             Fibrinogen C-terminal domain-like

   1fzc at  381 B   2.3        BLOOD COAGULATION             Fibrinogen C-terminal domain-like

   1alv at  223 B   1.9        CALCIUM BINDING               EF Hand-like / EF-hand modules in multidomain proteins

   1alv at  223 A   1.9        CALCIUM BINDING               EF Hand-like / EF-hand modules in multidomain proteins

   2por at   93      1.8       INTEGRAL MEMBRANE PROTEIN PORIN  Transmembrane beta-barrelS/Porin



pc 1-13

            13 protein chains to be compared

start and end seq values for fi -10  11; for psi  -10  11
```

```
   relseq      meanfi          meanpsi       number

    -10        -67( 10)        -35( 16)        13        a

     -9        -71( 17)        -34( 19)        13        a

     -8        -61(  7)        -40(  7)        13        a

     -7        -63(  4)        -40(  7)        13        a

     -6        -63(  5)        -37(  7)        13        a

     -5        -63(  5)        -42(  5)        13        a

     -4        -59(  7)        -43(  6)        13        a

     -3        -59(  6)        -36( 13)        13        a

     -2        -63(  6)        -33(  8)        13        a

     -1        -92( 17)        -19( 24)        13        a

      0        -79(  6)         78(  7)        13        d

      1        -59(  8)        -41( 17)        13        a

      2        -87( 13)          3(  8)        13        k

      3         59( 10)         29( 10)        13        g

      4        -84(  7)         -1( 10)        13        k

      5         88(  8)          0(  8)        13        g

      6       -137(  7)        155( 11)        13        b

      7        -98( 11)        121(  9)        13        b

      8        -97( 19)        171(  6)        13        b

      9        -57(  7)        -40(  8)        13        a

     10        -62(  8)        -36(  7)        13        a

     11        -69(  7)        -35(  7)        13        a



pc 16-23

                8 protein chains to be compared

start and end seq values for fi -10  11; for psi  -10  11


   relseq      meanfi          meanpsi       number

    -10        -58(  8)        -40( 10)         6        a

     -9        -65(  5)        -38(  7)         6        a

     -8        -64(  5)        -38(  7)         6        a
```

```
     -7         -62(   6)         -44(   5)           6          a

     -6         -56(   9)         -40(   7)           6          a

     -5         -61(   4)         -32(  12)           6          a

     -4         -66(   4)         -32(   8)           6          a

     -3         -83(  11)           0(  66)           8          k

     -2         -96(  33)          76(   5)           8          d

     -1         -54(  10)         -39(  14)           8          a

      0         -89(  11)           0(   7)           8          k

      1          62(   9)          25(   7)           8          g

      2         -85(   5)           0(  11)           8          k

      3          93(  14)          -1(  14)           8          g

      4        -127(  25)         146(  15)           8          b

      5         -90(  25)         127(  12)           8          b

      6         -88(  28)         207(  67)           8          .

      7         -59(   4)         -37(   7)           8          a

      8         -64(   5)         -39(   6)           8          a

      9         -64(   5)         -39(   6)           8          a

     10         -66(   4)         -43(   5)           8          a

     11         -66(  11)         -39(   6)           8          a




pc 24,25,26

                3 protein chains to be compared

start and end seq values for fi -10   11; for psi  -10   11


   relseq      meanfi             meanpsi        number

    -10         -60(   2)         133(  10)           3          b

     -9          83(   4)           0(   8)           3          g

     -8         -82(   2)         146(  13)           3          b

     -7         -80(  11)         159(   6)           3          b

     -6         -80(   6)         143(   4)           3          b
```

```
 -5        -131(   6)         147(   4)            3          b
 -4        -134(   8)         165(   2)            3          b
 -3         -60(   6)         -24(  10)            3          a
 -2        -101(   8)          14(   2)            3          k
 -1        -158(   3)          87(   5)            3          b
  0         -84(   1)         118(   4)            3          b   **
  1        -134(   2)          27(   5)            3          k
  2         -82(   6)         183(   6)            3          b   **
  3         -98(   6)          14(   8)            3          k
  4        -104(   2)         138(  20)            3          b
  5         -76(  17)           6(  69)            3          k
  6         -98(  15)        -166(  18)            3          b
  7        -117(   6)          55(  77)            3          d
  8         -59(   6)           6(  96)            3          k
  9         -57(   3)         -38(   9)            3          a
 10         -69(   5)         -33(  10)            3          a
 11         -72(  10)           0(  70)            3          k
```

```
Ca D N 2
cadn2 sequences , and, added for comparison
                        protein chains 16,16 18 from cadd2   (were 9, 21, 25)


  amino-acid sequences ****
   start of selected part of cngroup        |
   relseq            -10        -5          0          5        10        15
 1  1cdl from   48 B:  L Q D M I N E V D A D G N G T I D F P E F L
 2  1cdl from   85 A:  I R E A F R V F D K D G N G Y I S A A E L R
 3  1g4y from   48 R:  L Q D M I N E V D A D G N G T I D F P E F L
 4  1rec from  102  :  L E W A F S L Y D V D G N G T I S K N E V L
 5  1vrk from   48 A:  L Q D M I N E V D A D G N G T I D F P E F L
 6  2sas from    9  :   K I K F T F D F F L D M N H D G S I Q D N D
 7  1a2x from  129 A:  E E I E S L M K D G D K N N D G R I D F D E
 8  1a2x from   93 A:  E E L A E C F R I F D R N A D G Y I D A E E
 9  2scp from   94 B:  G P L P L F F R A V D T N E D N N I S R D E
10  2sas from   60  :  D E W R D L K G R A D I N K D D V V S W E E
11  2scp from   94 A:  G P L P L F F R A V D T N E D N N I S R D E
12  2scp from  128 B:  T M A P A S F D A I D T N N D G L L S L E E
13  1gca from  124  :  A K H W Q A N Q G W D L N K D G K I Q Y V L
14  1i8a from   71 A:  W E Q D S V E I F I D E N N H K T G Y Y E D
15  1b0p from  973 B:  G L D H V L A S G E D V N V F V M D T E V Y
16  2pvb from   41 A:  D D V K K A F Y V I D Q D K S G F I E E D E
17  2pvb from   82 A:  T K A F L A D G D K D G D G M I G V D E F A
18  3fib from  308  :  N G M Q F S T W D N D N D K F E G N C A E Q


  aa types - acidic, etc **
   relseq            -10        -5          0          5        10        15
 1  1cdl from   48 B:  h p a h h p a h a s a g p g s h a h h a h h
 2  1cdl from   85 A:  h b a s h b h h a b a g p g h h s s s a h b
 3  1g4y from   48 R:  h p a h h p a h a s a g p g s h a h h a h h
 4  1rec from  102  :  h a h s h s h h a h a g p g s h s b p a h h
```

```
  5  1vrk from   48 A:  h p a h h p a h a s a g p g s h a h h a h h

  6  2sas from    9  :  b h b h s h a h h h a h p p a g s h p a p a

  7  1a2x from  129 A:  a a h a s h h b a g a b p p a g b h a h a a

  8  1a2x from   93 A:  a a h s a h h b h h a b p s a g h h a s a a

  9  2scp from   94 B:  g h h h h h h b s h a s p a a p p h s b a a

 10  2sas from   60  :  a a h b a h b g b s a h p b a a h h s h a a

 11  2scp from   94 A:  g h h h h h h b s h a s p a a p p h s b a a

 12  2scp from  128 B:  s h s h s s h a s h a s p p a g h h s h a a

 13  1gca from  124  :  s b p h p s p p g h a h p b a g b h p h h h

 14  1i8a from   71 A:  h a p a s h a h h h a a p p p b s g h h a a

 15  1b0p from  973 B:  g h a p h h s s g a a h p h h h h a s a h h

 16  2pvb from   41 A:  a a h b b s h h h h a p a b s g h h a a a a

 17  2pvb from   82 A:  s b s h h s a g a b a g a g h h g h a a h s

 18  3fib from  308  :  p g h p h s s h a p a p a b h a g p h s a p


   conformation sequences **

    relseq               -10        -5         0         5        10        15

  1  1cdl from   48 B:  a a k a a a a a b a k g a g b b b a a a a a

  2  1cdl from   85 A:  a a a a a k a a d a k g k g b b b a a a a a

  3  1g4y from   48 R:  a a a a a a a d b a k g k g b b b a a a k a

  4  1rec from  102  :  a a a a a a a a d a k g k g b b b a a a a a

  5  1vrk from   48 A:  a a a a a a a k d a k g k g b b b a a a a a

  6  2sas from    9  :  a a a a a a a a k . d a k g k g b b b a a a

  7  1a2x from  129 A:  a a a a a a a a a a b a k g k g b b b a a a

  8  1a2x from   93 A:  a a a a a a a a a a b a k g k g b b b a a a

  9  2scp from   94 B:  a a a a a a a a a a d a k . k g b b b a a a

 10  2sas from   60  :  a a a a a a a k a a b a k g k g b b b a a a

 11  2scp from   94 A:  a k a a a a a a a a d a k g k g b b b a a a

 12  2scp from  128 B:  k k a a a a a a a a d k k . k g b b b a a a

 13  1gca from  124  :  a a a a a k d a k k d a k g k g b b b b b b

 14  1i8a from   71 A:  k k b b . b b b b b b a k g k g b k b b b a

 15  1b0p from  973 B:  a a a a a a a a g b b b b b b b b b k g . b

 16  2pvb from   41 A:  a a a a a a a a a a d a k g k g b b b a a a
```

```
17  2pvb from   82 A:  a a a a a a a a d k k g k g b b b a a a a a

18  3fib from  308  :  b g b b b b b a k b b k b k b a . b a a a a


  **********  protein names (pdb header) *******

  1  1cdl at   58 B   2.2       CALCIUM-BINDING PROTEIN

  2  1cdl at   95 A   2.2       CALCIUM-BINDING PROTEIN

  3  1g4y at   58 R   1.6       SIGNALING PROTEIN

  4  1rec at  112     1.9       CALCIUM-BINDING PROTEIN

  5  1vrk at   58 A   1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT

  6  2sas at   19      2.4       CALCIUM-BINDING PROTEIN

  7  1a2x at  139 A   2.3       COMPLEX (SKELETAL MUSCLE/MUSCLE PROT

  8  1a2x at  103 A   2.3       COMPLEX (SKELETAL MUSCLE/MUSCLE PROT

  9  2scp at  104 B   2.0       BINDING PROTEIN

 10  2sas at   70     2.4       CALCIUM-BINDING PROTEIN

 11  2scp at  104 A   2.0       BINDING PROTEIN

 12  2scp at  138 B   2.0       BINDING PROTEIN

 13  1gca at  134     1.7       GALACTOSE-BINDING PROTEIN

 14  1i8a at   81 A   1.90      HYDROLASE

 15  1b0p at  983 B   2.31      OXIDOREDUCTASE

 16  2pvb at   51 A   0.91      METAL BINDING PROTEIN

 17  2pvb at   92 A   0.91      METAL BINDING PROTEIN

 18  3fib at  318     2.1       BLOOD COAGULATION


  **********  cngps  ************

  1 2  DDNNOE   Ca   2   2   0   2   5  -1  -1 ......6 1cdl  0 0.6 2.2  CA     2 ASPB  56 .....b  |

  2 2  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  0 0.3 2.2  CA     3 ASPA  93 .b..b   |

  3 2  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1g4y  0 0.4 1.6  CA 1002 ASPR  56 ....b   |

  4 2  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1rec  0 0.2 1.9  CA  501 ASP  110 ....b   |     Z

  5 2  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.1 1.9  CA  152 ASPA  56 ....b   |     Z  2.

  6 1  DNDOD    Ca   2   2   2   5  -1  -1  -1 ..... 5 2sas  0 0.2 2.4  CA  186 ASP   19 ....b   |

  7 1  DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1a2x  1 0.4 2.3  CA  161 ASPA 139 ..b.b   |

  8 1  DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1a2x  2 0.2 2.3  CA  160 ASPA 103 ....b   |

  9 1  DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 2scp  0 0.3 2.0  CA  194 ASPB 104 ....b   |
```

```
10 1  DNDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 2sas  0 0.4 2.4  CA  187 ASP   70 ....b   |     Z
11 1  DNDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 2scp  0 0.1 2.0  CA  191 ASPA 104 ....b   |     Z
12 1  DNDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 2scp  0 0.1 2.0  CA  195 ASPB 138 ....b   |     Z
13 1  DNDOQE    Ca   2   2   2   2  63  -1  -1 ......6 1gca  0 0.1 1.7  CA    1 ASP  134 .....b  |
14 1  DNEDD     Ca   2   8   2   1  -1  -1  -1 ..... 6 1i8a  0 0.3 1.9  CA  190 ASPA  81 ...bb   |     Z  3.
15 1  DNOOO     Ca   2  71   3   2  -1  -1  -1 ..... 5 1b0p  1 0.4 2.3  CA 1238 ASPB 983 .....   |         1.
16 1  DDSOEE    Ca   2   2   2   2   3  -1  -1 ......6 2pvb  0 0.1 0.9  CA  110 ASPA  51 .....b  |
17 2  DDDOE     Ca   2   2   2   5  -1  -1  -1 ..... 6 2pvb  0 0.1 0.9  CA  111 ASPA  90 ....b   |     Z
18 1  DDOO      Ca   2   2   2  -1  -1  -1  -1 ....  6 3fib  0 0.2 2.1  CA  400 ASP  318 b...    |    ZZ
```

```
group 1:  p chains 6-14, agree within 10-30 deg, bak, and similar to
              group 1 of cadd2
            p chains 9 10 11 12 13 are within 8-15 deg, dak
                and are within 12 deg of p chain ch 16  i.e. group 1 of cadd2
              model 1gca at  134      1.7   (13)
              relseq -10 to 11:  9,10,11,12 are fairly similar, and 6,7,8


group 2:  p chains 1-5 agree 11-25 deg and same as 21 of cadd2   kgk
            model:  1g4y at   58 R   (3)
            chains are fairly similar to each other over relseq -10 to 11, and top 17
```

```
p chains 1-5
 relseq     meanfi          meanpsi       number
      0        -79( 14)          1( 10)          5        k
```

```
               1           59( 11)           29(  8)            5          g

               2         -85( 12)            0( 14)            5          k


        p chains 6-14


           relseq      meanfi             meanpsi        number

               0         -86( 17)           85( 14)            9          b

               1         -67( 17)         -34( 11)            9          a

               2         -94(  7)             8( 13)            9          k


          p chains 9-14

           relseq      meanfi             meanpsi        number

               0         -79(  3)           78(  6)            5          d

               1         -61(  6)         -31( 12)            5          a

               2         -98(  6)           17(  9)            5          k
```

```
Ca D O 2
cado2    11.6.02
  amino-acid sequences ****
   start of selected part of cngroup      |
   relseq              -10         -5          0          5         10         15
 1  2scp from    10 A:  T Y F N R I D F D K D G A I T R M D F E S M
 2  1cdl from    14 A:  E A F S L F D K D G D G T I T T K E L G T V
 3  1cdl from   123 A:  E M I R E A D I D G D G Q V N Y E E F V Q M
 4  1g4y from    14 R:  E A F S L F D K D G D G T I T T K E L G T V
 5  1vrk from    14 A:  E A F S L F D K D G D G T I T T K E L G T V
 6  2pvb from    84 A:  A F L A D G D K D G D G M I G V D E F A A M
 7  2sas from    13  :  T F D F F L D M N H D G S I Q D N D F E D M
 8  1a2x from    97 A:  E C F R I F D R N A D G Y I D A E E L A E I
 9  2sas from   109  :  F L F K G M D V S G D G I V D L E E F Q N Y
10  2por from   126  :  G A F S V A A S M S D G K V G E T S E D D A
11  1vrk from   123 A:  E M I R E A D V D G D G Q V N Y E E F V Q V
12  1gca from   128  :  Q A N Q G W D L N K D G K I Q Y V L L K G E
13  1sra from   251  :  R F F E T C D L D N D K Y I A L D E W A G C
14  1kap from   440 P:  A G S L A I D F S G D A H A D F A I N L I G
15  2scp from    98 B:  L F F R A V D T N E D N N I S R D E Y G I F
16  1ezm from   173  :  A A E F Y M R G K N D F L I G Y D I K K G S
17  1alv from   100 B:  R R L F A Q L A G D D M E V S A T E L M N I
18  2sas from    64  :  D L K G R A D I N K D D V V S W E E Y L A M
19  2scp from   132 B:  A S F D A I D T N N D G L L S L E E F V I A
20  1sra from   217  :  Q F G Q L D Q H P I D G Y L S H T E L A P L
21  1a2x from   133 A:  S L M K D G D K N N D G R I D F D E F L K M
22  1wdc from    15 C:  F E L F D F W D G R D G A V D A F K L G D V
23  1acc from   171  :  _ _ _ _ _ P D R D N D G I P D S L E V E G Y
24  1ajj from    15  :  G E C I H S S W R C D G G P D C K D K S D E
25  1i8a from     2 A:  A T A K Y G T P V I D G E I D E I W N T T E

26  1fzc from   373 B:  M F F S T Y D R D N D G W L T S D P R K Q C
```

```
27  1fzc from  310 C:  M Q F S T W D N D N D K F E G N C A E Q D G

28  3fib from  310  :  M Q F S T W D N D N D K F E G N C A E Q D G

29  1clc from  513  :  M N P H D R R S G A D G I W E P W P G Y L V

30  1nls from    0  :  _ _ D T I V A V E L D T Y P N T D I G D P S

31  1qho from   11 A:  I Y Q I I I D R F Y D G D T T N N N P A K S

32  1d2s from   40 A:  T S S S F E V R T W D P E G V I F Y G D T N

33  1g0h from  371 B:  D N S S E W T V V I D P I D G S F N F I N G

34  1g0h from   71 A:  D N S S E W T V V I D P I D G S F N F I N G

35  1i8a from   50 A:  N Y L Y V L A I V K D P V L N K D N S N P W

36  1fjs from   60 A:  Y A K R F K V R V G D R N T E Q E E G G E A

37  1cvr from   93 A:  D Q V Y G Q I V G N D H Y N E V F I G R F S

38  1kap from  275 P:  S S S K L V F S V W D A G G N D T L D F S G

39  1aru from  192  :  L N S A I F R S P L D S T P Q V F D T Q F Y


  aa types - acidic, etc **

   relseq              -10         -5          0          5         10         15

 1  2scp from   10 A:  s h h p b h a h a b a g s h s b h a h a s h

 2  1cdl from   14 A:  a s h s h h a b a g a g s h s s b a h g s h

 3  1cdl from  123 A:  a h h b a s a h a g a g p h p h a a h h p h

 4  1g4y from   14 R:  a s h s h h a b a g a g s h s s b a h g s h

 5  1vrk from   14 A:  a s h s h h a b a g a g s h s s b a h g s h

 6  2pvb from   84 A:  s h h s a g a b a g a g h h g h a a h s s h

 7  2sas from   13  :  s h a h h h a h p p a g s h p a p a h a a h

 8  1a2x from   97 A:  a h h b h h a b p s a g h h a s a a h s a h

 9  2sas from  109  :  h h h b g h a h s g a g h h a h a a h p p h

10  2por from  126  :  g s h s h s s s h s a g b h g a s s a a a s

11  1vrk from  123 A:  a h h b a s a h a g a g p h p h a a h h p h

12  1gca from  128  :  p s p p g h a h p b a g b h p h h h h b g a

13  1sra from  251  :  b h h a s h a h a p a b h h s h a a h s g h

14  1kap from  440 P:  s g s h s h a h s g a s p s a h s h p h h g

15  2scp from   98 B:  h h h b s h a s p a a p p h s b a a h g h h

16  1ezm from  173  :  s s a h h h b g b p a h h h g h a h b b g s

17  1alv from  100 B:  b b h h s p h s g a a h a h s s s a h h p h
```

```
18  2sas from   64  :  a h b g b s a h p b a a h h s h a a h h s h

19  2scp from  132 B:  s s h a s h a s p p a g h h s h a a h h h s

20  1sra from  217  :  p h g p h a p p h h a g h h s p s a h s h h

21  1a2x from  133 A:  s h h b a g a b p p a g b h a h a a h h b h

22  1wdc from   15 C:  h a h h a h h a g b a g s h a s h b h g a h

23  1acc from  171  :  _ _ _ _ _ h a b a p a g h h a s h a h a g h

24  1ajj from   15  :  g a h h p s s h b h a g g h a h b a b s a a

25  1i8a from    2 A:  s s s b h g s h h h a g a h a a h h p s s a


26  1fzc from  373 B:  h h h s s h a b a p a g h h s s a h b b p h

27  1fzc from  310 C:  h p h s s h a p a p a b h a g p h s a p a g

28  3fib from  310  :  h p h s s h a p a p a b h a g p h s a p a g

29  1clc from  513  :  h p h p a b b s g s a g h h a h h h g h h h

30  1nls from    0  :  _ _ a s h h s h a h a s h h p s a h g a h s

31  1qho from   11 A:  h h p h h h a b h h a g a s s p p p h s b s

32  1d2s from   40 A:  s s s s h a h b s h a h a g h h h h g a s p

33  1g0h from  371 B:  a p s s a h s h h h a h h a g s h p h h p g

34  1g0h from   71 A:  a p s s a h s h h h a h h a g s h p h h p g

35  1i8a from   50 A:  p h h h h h s h h b a h h h p b a p s p h h

36  1fjs from   60 A:  h s b b h b h b h g a b p s a p a a g g a s

37  1cvr from   93 A:  a p h h g p h h g p a p h p a h h h g b h s

38  1kap from  275 P:  s s s b h h h s h h a s g g p a s h a h s g

39  1aru from  192  :  h p s s h h b s h h a s s h p h h a s p h h


   conformation sequences **

     relseq                -10        -5         0         5        10        15

 1  2scp from   10 A:  a a a a a k d a k g k g b b b a a a a a a a

 2  1cdl from   14 A:  a a a a a a d a k g k g b b b a a a a a a a

 3  1cdl from  123 A:  a a a a a k d a k g k g b b b a a a a a a a

 4  1g4y from   14 R:  a a a a k a b a k g k g b b b a a a a a a a

 5  1vrk from   14 A:  a a a a a a d a k g k g b b b a a a a a a a

 6  2pvb from   84 A:  a a a a a a d k k g k g b b b a a a a a a a

 7  2sas from   13  :  a a a a k . d a k g k g b b b a a a a a a a
```

```
 8  1a2x from   97 A:  a a a a a a b a k g k g b b b a a a a a a a
 9  2sas from  109  :  a a a k k a d a k g k g b b b a a a a a a a
10  2por from  126  :  . k b b b b b b * b k g b b . d b k b b b b    * near k
11  1vrk from  123 A:  a a a a a k b a k g k g b b b a a a a a a a
12  1gca from  128  :  a k d a k k d a k g k g b b b b b b b b . b
13  1sra from  251  :  a a a a a k d a k g k g b b b a a a a a a a
14  1kap from  440 P:  b b . b b b b a a g k g b b a b b b b b b j
15  2scp from   98 B:  a a a a a a d a k * k g b b b a a a a a a a    * near g
16  1ezm from  173  :  a a a a a a k g b b k g b b g a a k b b j a
17  1alv from  100 B:  a a a a a a a a j a k g b b b a a a a a a a
18  2sas from   64  :  a a a k a a b a k g k g b b b a a a a a a a
19  2scp from  132 B:  a a a a a a d k k * k g b b b a a a a a a a    * near g
20  1sra from  217  :  a a a a a b k d b b * g b b b a a k k a a k    * far from g
21  1a2x from  133 A:  a a a a a a b a k g k g b b b a a a a a a a
22  1wdc from   15 C:  a a a a a a k k * b - g b b b k k k a a a a    * far from g
23  1acc from  171  :  . . . . . b d a k g k g b b a a a a a a . b     - v near k
24  1ajj from   15  :  g b b b b a k k b g k g b b a b a k g a k k
25  1i8a from    2 A:  b b b b b . b b b b * g b b b a a a k k b b    * near k

26  1fzc from  373 B:  b b b b b k k b b k b k b b b a b a a d b a
27  1fzc from  310 C:  b b b b b a k . b d b k b a . b a a a a k g
28  3fib from  310  :  b b b b b a k b b k b k b a . b a a a a a g
29  1clc from  513  :  a d b b b a a a k k b k b a b b b b g b b b
30  1nls from    0  :  . . b b b b b b b b b k b b b a k k g b b d
31  1qho from   11 A:  b b b b b a k k a b b k b k k k a d b a a d
32  1d2s from   40 A:  d b b b . b b b . k b k b . b a b b b b k .
33  1g0h from  371 B:  b k g b k b b b b b b a b b g a a a a k k g
34  1g0h from   71 A:  b k g b k b b b b b b a b b g a a a a a k g
35  1i8a from   50 A:  a b b b b b b b b b . k b b b d a b a b a k
36  1fjs from   60 A:  d b a b b b b b b g b k b a d . b a g k b b
37  1cvr from   93 A:  a a a a k b a b g g b k b b a b b b b b b b
38  1kap from  275 P:  k k b b b b . b b b b k . j b b b b b d a k
39  1aru from  192  :  a b k a d b g b b k b k d k k b b b k a a a
```

```
 **********  protein names (pdb header) *******
 1  2scp at   20 A   2.0       BINDING PROTEIN
 2  1cdl at   24 A   2.2       CALCIUM-BINDING PROTEIN
 3  1cdl at  133 A   2.2       CALCIUM-BINDING PROTEIN
 4  1g4y at   24 R   1.6       SIGNALING PROTEIN
 5  1vrk at   24 A   1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT
 6  2pvb at   94 A   0.91      METAL BINDING PROTEIN
 7  2sas at   23      2.4       CALCIUM-BINDING PROTEIN
 8  1a2x at  107 A   2.3       COMPLEX (SKELETAL MUSCLE/MUSCLE PROT
 9  2sas at  119      2.4       CALCIUM-BINDING PROTEIN
10  2por at  136      1.8       INTEGRAL MEMBRANE PROTEIN PORIN
11  1vrk at  133 A   1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT
12  1gca at  138      1.7       GALACTOSE-BINDING PROTEIN
13  1sra at  261      2.0       CALCIUM-BINDING PROTEIN
14  1kap at  450 P   1.64      ZINC METALLOPROTEASE
15  2scp at  108 B   2.0       BINDING PROTEIN
16  1ezm at  183      1.5       HYDROLASE
17  1alv at  110 B   1.9       CALCIUM BINDING
18  2sas at   74      2.4       CALCIUM-BINDING PROTEIN
19  2scp at  142 B   2.0       BINDING PROTEIN
20  1sra at  227      2.0       CALCIUM-BINDING PROTEIN
21  1a2x at  143 A   2.3       COMPLEX (SKELETAL MUSCLE/MUSCLE PROT
22  1wdc at   25 C   2.0       MUSCLE PROTEIN
23  1acc at  181      2.1       TOXIN
24  1ajj at   25      1.7       RECEPTOR
25  1i8a at   12 A   1.90      HYDROLASE
26  1fzc at  383 B   2.3       BLOOD COAGULATION
27  1fzc at  320 C   2.3       BLOOD COAGULATION
28  3fib at  320      2.1       BLOOD COAGULATION
29  1clc at  523      1.9       GLYCOSYL HYDROLASE
30  1nls at   10      0.94      AGGLUTININ
31  1qho at   21 A   1.70      HYDROLASE
```

```
32  1d2s  at    50  A    1.55      TRANSPORT  PROTEIN
33  1g0h  at   381  B    2.30      HYDROLASE
34  1g0h  at    81  A    2.30      HYDROLASE
35  1i8a  at    60  A    1.90      HYDROLASE
36  1fjs  at    70  A    1.92      BLOOD  CLOTTING
37  1cvr  at   103  A    2.00      HYDROLASE
38  1kap  at   285  P    1.64      ZINC  METALLOPROTEASE
39  1aru  at   202       1.6       PEROXIDASE  (DONOR:H2O2  OXIDOREDUCTAS


  **********   cngps   ************
  1  3   DDDOD     Ca    2    2    2    5   -1   -1   -1  .....  6  2scp   0  0.1  2.0   CA   190  ASPA   16  ....b    |     Z
  2  3   DDDOE     Ca    2    2    2    5   -1   -1   -1  .....  5  1cdl   0  0.4  2.2   CA     1  ASPA   20  .b..b    |
  3  3   DDDOE     Ca    2    2    2    5   -1   -1   -1  .....  5  1cdl   1  0.6  2.2   CA     4  ASPA  129  .bb..    |
  4  3   DDDOE     Ca    2    2    2    5   -1   -1   -1  .....  6  1g4y   0  0.2  1.6   CA  1001  ASPR   20  ....b    |     Z
  5  3   DDDOE     Ca    2    2    2    5   -1   -1   -1  .....  6  1vrk   0  0.1  1.9   CA   151  ASPA   20  ....b    |     Z   2.
  6  3   DDDOE     Ca    2    2    2    5   -1   -1   -1  .....  6  2pvb   0  0.1  0.9   CA   111  ASPA   90  ....b    |     Z
  7  3   DNDOD     Ca    2    2    2    5   -1   -1   -1  .....  5  2sas   0  0.2  2.4   CA   186  ASP    19  ....b    |
  8  3   DNDOE     Ca    2    2    2    5   -1   -1   -1  .....  5  1a2x   2  0.2  2.3   CA   160  ASPA  103  ....b    |
  9  3   DSDOE     Ca    2    2    2    5   -1   -1   -1  .....  5  2sas   0  0.2  2.4   CA   188  ASP   115  ....b    |
 10  2   NDOO      Ca   20    2    2   -1   -1   -1   -1  ....   5  2por   0  0.2  1.8   CA   304  ASN   116  .b..     |    Z
 11  3   DDDOE     Ca    2    2    2    5   -1   -1   -1  .....  6  1vrk   0  0.2  1.9   CA   154  ASPA  129  ....b    |     Z   2.
 12  3   DNDOQE    Ca    2    2    2    2   63   -1   -1  ......6  1gca   0  0.1  1.7   CA     1  ASP   134  .....b   |
 13  3   DDDOE     Ca    2    2    2    5   -1   -1   -1  .....  6  1sra   0  0.2  2.0   CA   302  ASP   257  ....b    |     Z
 14  3   DSDOD     Ca    2    2    2    2   -1   -1   -1  ... .  6  1kap   0  0.3  1.6   CA   621  ASPP  446  ....b    |     Z   3.
 15  3   DNDOE     Ca    2    2    2    5   -1   -1   -1  .....  5  2scp   0  0.3  2.0   CA   194  ASPB  104  ....b    |
 16  4   DEEDO     Ca   36    3    8    2   -1   -1   -1  .....  6  1ezm   0  0.3  1.5   CA   400  ASP   136  ..b..    |     Z   3.
 17  2   ODOE      Ca    3    2    5   -1   -1   -1   -1  ....   6  1alv   0  0.1  1.9   CA     5  ALAB  107  ...b     |    ZZ   3.
 18  3   DNDOE     Ca    2    2    2    5   -1   -1   -1  .....  6  2sas   0  0.4  2.4   CA   187  ASP    70  ....b    |     Z
 19  3   DNDOE     Ca    2    2    2    5   -1   -1   -1  .....  6  2scp   0  0.1  2.0   CA   195  ASPB  138  ....b    |     Z
 20  3   DODOE     Ca    3    2    2    5   -1   -1   -1  .....  6  1sra   0  0.2  2.0   CA   301  ASP   222  ....b    |     Z
 21  3   DNDOE     Ca    2    2    2    5   -1   -1   -1  .....  5  1a2x   1  0.4  2.3   CA   161  ASPA  139  ..b.b    |
 22  5   ODDODO    Ca    0    3    1    2    2   -1   -1  ......7  1wdc   0  0.2  2.0   CA   501  ASPC   19  ......   |      Z
 23  3   DDDOE     Ca    2    2    2    5   -1   -1   -1  .....  6  1acc   1  0.2  2.1   CA   800  ASP   177  .....    |     Z
```

```
24 2   ODODDE    Ca    3    2    2    6    1   -1   -1 ......6 1ajj   0 0.1 1.7   CA    73 TRP    22 ......   |
25 2   ODODE     Ca    2    2    2 114   -1   -1   -1 ..... 6 1i8a   0 0.2 1.9   CA   191 VALA   10 ....b    |     Z   3.
26 2   DDO        Ca    2    2   -1   -1   -1   -1   -1 ...    5 1fzc   0 0.2 2.3   CA     2 ASPB 381 b..       |   ZZ
27 2   DDOO      Ca    2    2    2   -1   -1   -1   -1 ....   4 1fzc   0 0.2 2.3   CA     1 ASPC 318 b...      |
28 2   DDOO      Ca    2    2    2   -1   -1   -1   -1 ....   6 3fib   0 0.2 2.1   CA   400 ASP   318 b...      |    ZZ
29 2   ODO        Ca    3    2   -1   -1   -1   -1   -1 ...    6 1clc   0 0.1 1.9   CA   593 SER   520 ...       |   ZZZ   3.
30 1   DOND      Ca    2    2    5   -1   -1   -1   -1 ....   6 1nls   0 0.1 0.9   CA   240 ASP    10 b...      |    ZZ
31 1   DONNOD    Ca    2    3    1   21    2   -1   -1 ......7 1qho   0 0.2 1.7   CA   698 ASPA   21 ......   |      Z 3.
32 1   DOO        Ca    2 108   -1   -1   -1   -1   -1 ...    6 1d2s   0 0.1 1.5   CA   401 ASPA   50 b..       |   ZZZ
33 2   EDO        Ca   16    2   -1   -1   -1   -1   -1 ...    4 1g0h   1 0.4 2.3   CA   590 GLUB 365 .b.       |   U      3.
34 2   EDO        Ca   16    2   -1   -1   -1   -1   -1 ...    6 1g0h   0 0.2 2.3   CA   291 GLUA   65 b..       |   UZZ   3.
35 1   DODDO     Ca    2   12   80    1   -1   -1   -1 ..... 6 1i8a   0 0.3 1.9   CA   192 ASPA   60 b.b..    |     Z   3.
36 1   DOOEE     Ca    2    3    2    3   -1   -1   -1 ..... 6 1fjs   0 0.3 1.9   CA   507 ASPA   70 .....    |     Z   3.
37 2   ODOE      Ca    3    2    2   -1   -1   -1   -1 ....   6 1cvr   0 0.2 2.0   CA   477 VALA 100 ...b      |    ZZ   3.
38 4   OOTDOD    Ca    2    2   28    2    3   -1   -1 ......6 1kap   1 0.1 1.6   CA   614 ARGP 253 ...b..   |         3.
39 3   OSDOTOD   Ca    0   17    2    0    3    2   -1 ......7 1aru   0 0.2 1.6   CA   347 SER   185 ..b.... |         1.



over donor pair 0-2,0-1

 1-25 are similar, within 30-40 deg    -    kgb

  1-10 are within 5-14 deg                      model:2pvb at    94 A    0.9    (6)

  14-19 are within 12 deg of each other    model:1ezm at   183       1.5    (16)

  23-24 are within 17 deg of each other    model:1ajj at    25       1.7    (24)


 26-39 are similar, within 50-60 deg of each other, 80+ deg from above - bkb or bab

  26-29 are very close  (10 deg)               model:3fib at   320       2.1     (28)

  30-32 are very close (13 deg)                model:1nls at    10       0.94    (30)

  33-34 are very close (10 deg) - probably identical (A&B), sloppy crystallography

     39 fairly close to 34 35 (16 deg)


how much of this is just as a part of whole chain -10 to 11 ?


          donor pair ( 0-2,0-1)                      whole chain
```

```
   1-10 are within 5-14 deg                     10 very different, also 12 14

   14-19 are within 12 deg of each other:   16 different

   23-25 are within 14 deg of each other    20, 22-25 fairly different


  26-39 are similar, within 50-60 deg of each other, 80+ deg from above

   26-29 are very close   (8 deg)              all different (50-80 deg)

   30-32 are very close (11 deg)               all different (80 deg)


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


p chains 1-25

   relseq      meanfi             meanpsi        number

      0        -100( 22)            6( 10)           25          k

      1           80( 15)            5( 17)           25          g

      2        -128( 17)         149( 18)           25          b


p chains 1-10

  relseq      meanfi             meanpsi        number

      0          -89(  9)            3(  7)           10          k

      1           86(  6)            5(  9)           10          g

      2        -140(  5)          151( 13)           10          b


p chains 14-19

   relseq      meanfi             meanpsi        number

      0        -103(  7)           12(  9)            6          k

      1           60(  6)           21(  6)            6          g

      2        -130(  7)          138( 11)            6          b


p chains 23-24

   relseq      meanfi             meanpsi        number

      0          -96( 15)           -6( 15)            2          k

      1          108(  8)          -33( 15)            2          g

      2          -88(  1)          130(  2)            2          b
```

```
p chains 26-39
   relseq      meanfi             meanpsi        number
       0        -110( 31)        -214( 38)          13         b
       1          -89( 20)          -2( 19)          13         k
       2        -126( 27)         125( 27)          13         b

p chains 26-29
   relseq      meanfi             meanpsi        number
       0          -83(  6)        -175(  5)            4         b
       1        -100(  6)           17(  8)            4         k
       2        -103(  2)         138( 17)            4         b

p chains 30-32
   relseq      meanfi             meanpsi        number
       0          -91(  6)         106( 13)            3         b
       1          -81(  8)         -11(  2)            3         k
       2        -138(  8)         124( 22)            3         b

- - - - - - - - - - - - - - - - - - - - -- - - -
side chains  of 1-25
     all but 3 have chi1  40 to 82, chi2  -52 to 83
                3 has          68                   156   (probably other OD)
side chains of 26-39   widely scattered, but
     for prot chains 26-29   chi1 63 to 73,      chi2   7 to 37
                          30-32   chi1 -177 to -171, chi2  -18 to -1
```

```
Ca NO 2     and compare with Ca DD2 - included as protein chains 14 and 15



  amino-acid sequences ****

   start of selected part of cngroup         |

   relseq             -10         -5          0          5         10         15

 1  1cdl from   50 B: D M I N E V D A D G N G T I D F P E F L T M

 2  1cdl from   87 A: E A F R V F D K D G N G Y I S A A E L R H V

 3  1g4y from   50 R: D M I N E V D A D G N G T I D F P E F L T M

 4  1rec from  104  : W A F S L Y D V D G N G T I S K N E V L E I

 5  1vrk from   50 A: D M I N E V D A D G N G T I D F P E F L N L

 6  1bfd from  445  : N I P T I F V I M N N G T Y G A L R W F A G

 7  1trk from  177 A: L G N L I A I Y D D N K I T I D G A T S I S

 8  2cbl from  223 A: A L K S T I D L T C N D Y I S V F E F D I F

 9  1clc from  229  : P V A L E I P E K N N S I P D F L D E L K Y

10  1cse from   67 E: H V A G T V A A L D N T T G V L G V A P S V

11  1gci from   67  : H V A G T I A A L N N S I G V L G V A P S A

12  1scj from   67 A: H V A G T I A A L N N S I G V L G V S P S A

13  2sic from   67 E: H V A G T V A A L N N S I G V L G V A P S A

14  2pvb from   84 A: A F L A D G D K D G D G M I G V D E F A A M

15  1nls from    0  : _ _ D T I V A V E L D T Y P N T D I G D P S


  aa types - acidic, etc **

   relseq             -10         -5          0          5         10         15

 1  1cdl from   50 B: a h h p a h a s a g p g s h a h h a h h s h

 2  1cdl from   87 A: a s h b h h a b a g p g h h s s s a h b p h

 3  1g4y from   50 R: a h h p a h a s a g p g s h a h h a h h s h

 4  1rec from  104  : h s h s h h a h a g p g s h s b p a h h a h

 5  1vrk from   50 A: a h h p a h a s a g p g s h a h h a h h p h

 6  1bfd from  445  : p h h s h h h h h p p g s h g s h b h h s g

 7  1trk from  177 A: h g p h h s h h a a p b h s h a g s s s h s

 8  2cbl from  223 A: s h b s s h a h s h p a h h s h h a h a h h
```

```
 9  1clc from  229  :  h h s h a h h a b p p s h h a h h a a h b h
10  1cse from   67 E:  p h s g s h s s h a p s s g h h g h s h s h
11  1gci from   67  :  p h s g s h s s h p p s h g h h g h s h s s
12  1scj from   67 A:  p h s g s h s s h p p s h g h h g h s h s s
13  2sic from   67 E:  p h s g s h s s h p p s h g h h g h s h s s
14  2pvb from   84 A:  s h h s a g a b a g a g h h g h a a h s s h
15  1nls from    0  :  _ _ a s h h s h a h a s h h p s a h g a h s


   conformation sequences **
    relseq             -10        -5         0         5        10        15
 1  1cdl from   50 B:  k a a a a a b a k g a g b b b a a a a a a a
 2  1cdl from   87 A:  a a a k a a d a k g k g b b b a a a a a a a
 3  1g4y from   50 R:  a a a a a d b a k g k g b b b a a a k a a a
 4  1rec from  104  :  a a a a a a d a k g k g b b b a a a a a a a
 5  1vrk from   50 A:  a a a a a k d a k g k g b b b a a a a a a a
 6  1bfd from  445  :  g b d b b b b b b b k g b d a a a a a a a a
 7  1trk from  177 A:  b a k b b b b b b d k g b b a k j b a a a .
 8  2cbl from  223 A:  a a a a a a d a k g k g b b b a a a a a a a
 9  1clc from  229  :  k b b b d b a k k k b k b b a a a a a a a a
10  1cse from   67 E:  a a a a a a k b b b . k b j b b b a b k k b
11  1gci from   67  :  a a a a a a . b b b . k b j b b b a d k k b
12  1scj from   67 A:  a a a a a a d b b b . k b . b b b a b k k b
13  2sic from   67 E:  a a a a a a d b b b . d b j b b b a d a k b
14  2pvb from   84 A:  a a a a a a d k k g k g b b b a a a a a a a
15  1nls from    0  :  . . b b b b b b b b b k b b b a k k g b b d


  **********  protein names (pdb header) *******
 1  1cdl at   60 B   2.2       CALCIUM-BINDING PROTEIN
 2  1cdl at   97 A   2.2       CALCIUM-BINDING PROTEIN
 3  1g4y at   60 R   1.6       SIGNALING PROTEIN
 4  1rec at  114     1.9       CALCIUM-BINDING PROTEIN
 5  1vrk at   60 A   1.9       COMPLEX(CALCIUM-BINDING PROTEIN/PEPT
 6  1bfd at  455     1.6       LYASE
```

```
  7  1trk at  187 A   2.0       TRANSFERASE(KETONE RESIDUES)
  8  2cbl at  233 A   2.1       COMPLEX (PROTO-ONCOGENE/PEPTIDE)
  9  1clc at  239      1.9       GLYCOSYL HYDROLASE
 10  1cse at   77 E   1.2       COMPLEX(SERINE PROTEINASE-INHIBITOR)
 11  1gci at   77      0.78      SUBTILISIN FROM BACILLUS LENTUS
 12  1scj at   77 A   2.0       HYDROLASE
 13  2sic at   77 E   1.8       COMPLEX (PROTEINASE/INHIBITOR)
 14  2pvb at   94 A   0.91      METAL BINDING PROTEIN
 15  1nls at   10      0.94      AGGLUTININ


  **********  cngps  ************
  1 4  DDNNOE   Ca   2   2   0   2   5  -1  -1 ......6 1cdl  0 0.6 2.2  CA     2 ASPB  56 .....b  |
  2 3  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  0 0.3 2.2  CA     3 ASPA  93 .b..b   |
  3 3  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1g4y  0 0.4 1.6  CA 1002 ASPR  56 ....b   |
  4 3  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1rec  0 0.2 1.9  CA  501 ASP  110 ....b   |    Z
  5 3  DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.1 1.9  CA  152 ASPA  56 ....b   |    Z  2.
  6 2  DNO       Ca  27   2  -1  -1  -1  -1  -1 ...   6 1bfd  0 0.2 1.6  CA  529 ASP  428 ...      |  UUZ  4.
  7 2  DNO       Ca  30   2  -1  -1  -1  -1  -1 ...   6 1trk  0 0.2 2.0  CA  681 ASPA 157 ...      |  UUZ  2.
  8 3  DTNOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 2cbl  0 0.1 2.1  CA  352 ASPA 229 ....b   |    Z
  9 2  ONODD    Ca   3   2   2   3  -1  -1  -1 ..... 6 1clc  0 0.1 1.9  CA  591 GLU  236 .....   |    Z  3.
 10 4  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1cse  0 0.1 1.2  CA  430 GLNE   2 .b....  |        3.
 11 4  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1gci  0 0.1 0.8  CA  277 GLN    2 .b....  |        3.
 12 4  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1scj  0 0.2 2.0  CA  381 GLNA   2 .b....  |        3.
 13 4  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 2sic  0 0.2 1.8  CA  501 GLNE   2 .b....  |        3.
 14 3  DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 2pvb  0 0.1 0.9  CA  111 ASPA  90 ....b   |    Z
 15 1  DOND     Ca   2   2   5  -1  -1  -1  -1 ....  6 1nls  0 0.1 0.9  CA  240 ASP   10 b...    |   ZZ




 p chains 1-8
   relseq     meanfi          meanpsi      number
      0        -96( 20)          4( 11)          9        k
      1         80( 14)          3( 10)          9        g
```

```
       2        -135( 10)         153( 15)           9         b


p chains 10-13
  relseq       meanfi            meanpsi        number
       0        -153(  2)        -149(  4)           4         .
       1        -108(  7)          19(  9)           4         k
       2        -135(  7)         164(  2)           4         b




group 1:  p chains 1-8, 8 agrees poorly, without it sds 9-16
              and this is close to cado2  no. 14
             p chains 1-5, and 8 are fairly similar over relseq -10 to 11, 6 7 dif
             model : 1bfd at  455      1.6


group 2:  p chains 10-13
               chain 9 is similar but difs of 30 + degs
             chains 10-13 are similar over relseq -10 to 11, 9 is very different
                (11-13 have same aa sequence)
             model : 1gci at   77      0.78   - no equiv in cadd2
                      conformation of relseq =0 is some way from alpha region


no chains like 15 of cadd2
```

```
Ca DO 1


   amino-acid sequences ****

     start of selected part of cngroup         |

     relseq               -10         -5          0          5         10         15

  1   2sns from    30  :  Q P M T F R L L L V D T P E T K H P K K G

  2   1oac from   668 A:  S K D N E S L D N T D A V V W M T T G T T

  3   1i8a from   144 A:  T V I G F N I Q V N D A N E K G Q R V G I

  4   1qho from    66 A:  V T T I W L S P V L D N L D T L A G T D N

  5   4sgb from   110 E:  N T T P K _ _ _ _ D D I T S A A _ _ _ N A

  6   1ava from   132 A:  D G T G N P D T G A D F G A A P D I D H L

  7   1byf from    97 B:  V Q I W S K Y N L L D D V G C G G A R R V

  8   1dx5 from   413 J:  Y I L D D G F I C T D I D E C E N G G F C

  9   1qq9 from    -7 A:  _ _ _ _ _ _ _ _ A P D I P L A N V K A H L

 10   1ga6 from   318 A:  A I S S T P S L V H D V K S G N N G Y G G

 11   1ee6 from    70 A:  D C T I T N V I W E D V G E D A L T L K S

 12   1oac from   523 A:  T H Q H I Y N F R L D L D V D G E N N S L

 13   1cb8 from   406 A:  W D K I P G I T S R D Y L T D R P L T K L

 14   1kit from   672  :  V F S N I S T G T V D A S I T R F E Q S D

 15   1edm from    54 C:  N P C L N G G S C K D D I N S Y E C W C P

 16   1edm from    37 C:  _ _ _ _ _ _ _ _ _ V D G D Q C E S N P C L

 17   1gcy from    -9 A:  _ _ _ _ _ _ _ _ _ _ D Q A G K S P N A V R

 18   1i76 from   144 A:  Q R D H G D N S P F D G P N G I L A H A F

 19   1wdc from    12 C:  K D V F E L F D F W D G R D G A V D A F K


   aa types - acidic, etc **

     relseq               -10         -5          0          5         10         15

  1   2sns from    30  :  p h h s h b h h h h a s h a s b p h b b g

  2   1oac from   668 A:  s b a p a s h a p s a s h h h h s s g s s

  3   1i8a from   144 A:  s h h g h p h p h p a s p a b g p b h g h

  4   1qho from    66 A:  h s s h h h s h h h a p h a s h s g s a p

  5   4sgb from   110 E:  p s s h b _ _ _ _ a a h s s s s _ _ _ p s
```

```
  6  1ava from  132 A:  a g s g p h a s g s a h g s s h a h a p h
  7  1byf from   97 B:  h p h h s b h p h h a a h g h g g s b b h
  8  1dx5 from  413 J:  h h h a a g h h h s a h a a h a p g g h h
  9  1qq9 from   -7 A:  _ _ _ _ _ _ _ _ s h a h h h s p h b s p h
 10  1ga6 from  318 A:  s h s s s h s h h p a h b s g p p g h g g
 11  1ee6 from   70 A:  a h s h s p h h h a a h g a a s h s h b s
 12  1oac from  523 A:  s p p p h h p h b h a h a h a g a p p s h
 13  1cb8 from  406 A:  h a b h h g h s s b a h h s a b h h s b h
 14  1kit from  672  :  h h s p h s s g s h a s s h s b h a p s a
 15  1edm from   54 C:  p h h h p g g s h b a a h p s h a h h h h
 16  1edm from   37 C:  _ _ _ _ _ _ _ _ _ h a g a p h a s p h h h
 17  1gcy from   -9 A:  _ _ _ _ _ _ _ _ _ _ a p s g b s h p s h b
 18  1i76 from  144 A:  p b a p g a p s h h a g h p g h h s p s h
 19  1wdc from   12 C:  b a h h a h h a h h a g b a g s h a s h b


   conformation sequences **
     relseq             -10        -5         0          5         10         15
  1  2sns from   30  :  . b . b b b b b g . b b b b . . . a . . .
  2  1oac from  668 A:  k a k g b b k b g b b b b b b b b . b b
  3  1i8a from  144 A:  b b b b b b b b b b b b k k g b b k b b
  4  1qho from   66 A:  b k b b b b b b k b b b b k b b d j b g b
  5  4sgb from  110 E:  b a d b . . . . . . b b a b b . . . . . b
  6  1ava from  132 A:  k g k b b b b k . b b b a k b b b b b k k
  7  1byf from   97 B:  b b b b a a k g b b b b b b k k b b b b b
  8  1dx5 from  413 J:  b b b b b g g b b b b b b a a a a j g a b
  9  1qq9 from   -7 A:  . . . . . . . . . b b b b a a a a a a a a
 10  1ga6 from  318 A:  a a a k d a k k b b b b a b j b d . b g g
 11  1ee6 from   70 A:  b b b b b g b b b k . b . a b a b b b a b
 12  1oac from  523 A:  b b b b b b b b b b b b b g j a b b b b
 13  1cb8 from  406 A:  k k d b b g b b . b b b a b b b b b a b b
 14  1kit from  672  :  a b b g b b k b b b b b b b b b b b b a k
 15  1edm from   54 C:  g k b b g g b b b b b b b g . b b b b b b
 16  1edm from   37 C:  . . . . . . . . . . b b b k k a k g k b b
```

```
17  1gcy from   -9 A:  . . . . . . . . . . . b b . b b k k g b d
18  1i76 from  144 A:  b . b . g k g d b b k j b j j b a . b b b
19  1wdc from   12 C:  a a a a a a a a a k k . b . g b b b k k k

 **********  protein names (pdb header) *******
 1  2sns at   40     1.5       HYDROLASE (PHOSPHORIC DIESTER)
 2  1oac at  678 A   2.0       OXIDOREDUCTASE
 3  1i8a at  154 A   1.90      HYDROLASE
 4  1qho at   76 A   1.70      HYDROLASE
 5  4sgb at  120 E   2.1       COMPLEX(SERINE PROTEINASE-INHIBITOR)
 6  1ava at  142 A   1.9       HYDROLASE INHIBITION
 7  1byf at  107 B   2.00      SUGAR BINDING PROTEIN
 8  1dx5 at  423 J   2.3       SERINE PROTEINASE
 9  1qq9 at    3 A   1.53      HYDROLASE
10  1ga6 at  328 A   1.00      HYDROLASE
11  1ee6 at   80 A   2.30      LYASE
12  1oac at  533 A   2.0       OXIDOREDUCTASE
13  1cb8 at  416 A   1.90      LYASE
14  1kit at  682     2.3       HYDROLASE
15  1edm at   64 C   1.5       COAGULATION FACTOR
16  1edm at   47 C   1.5       COAGULATION FACTOR
17  1gcy at    1 A   1.6       HYDROLASE
18  1i76 at  154 A   1.20      HYDROLASE
19  1wdc at   22 C   2.0       MUSCLE PROTEIN

 **********  cngps  ************
 1 2  DDO       Ca  19   1  -1  -1  -1  -1  -1 ...   4 2sns  0 0.3 1.5  CA    1 ASP   21 ...      |  U    3.
 2 4  DODDO     Ca   1   1 143   1  -1  -1  -1 ..... 6 1oac  0 0.1 2.0  CA  802 ASPA 533 .....   |    Z  1.
 3 4  DODDO     Ca   2  12  80   1  -1  -1  -1 ..... 6 1i8a  0 0.3 1.9  CA  192 ASPA  60 b.b..   |    Z  3.
 4 1  DODEE     Ca   1   2  22   1  -1  -1  -1 ..... 6 1qho  1 0.2 1.7  CA  696 ASPA  76 ....b   |    Z  3.
 5 2  ODOOY     Ca   0   1 122   0  -1  -1  -1 ..... 6 4sgb  0 0.1 2.1  CA    8 GLYE 120 .....   |    Z
 6 2  DDOOD     Ca  15   1   3   2  -1  -1  -1 ..... 6 1ava  1 0.2 1.9  CA  502 ASPA 127 .....   |    Z  3.
 7 3  ENDOD     Ca   3  18   1   0  -1  -1  -1 ..... 7 1byf  0 0.1 2.0  CA  201 GLUB  86 .....   |    UU
```

```
 8 1  DOENOO   Ca   1   2  13   1   3  -1  -1 ......7 1dx5  0 0.2 2.3  CA 1001 ASPJ 423 b.....  |     Z 3.
 9 1  DODD      Ca   1 258   4  -1  -1  -1  -1 ....  6 1qq9  0 0.2 1.5  CA  905 ASPA   3 ...b    |   ZZ  3.
10 1  DOOOD    Ca   1  15   2   2  -1  -1  -1 ..... 6 1ga6  0 0.1 1.0  CA  374 ASPA 328 .....   |    Z  3.
11 1  DOO       Ca   1  22  -1  -1  -1  -1  -1 ...   6 1ee6  0 0.3 2.3  CA  300 ASPA  80 ...      |  ZZZ  4.
12 1  DODDO    Ca   1   1 143   1  -1  -1  -1 ..... 6 1oac  0 0.1 2.0  CA  802 ASPA 533 .....   |    Z  1.
13 3  EDDO      Ca   2   9   1  -1  -1  -1  -1 ....  6 1cb8  0 0.2 1.9  CA 3000 GLUA 405 .b..    |   ZZ  4.
14 2  DDO       Ca  61   1  -1  -1  -1  -1  -1 ...   4 1kit  0 0.2 2.3  CA  803 ASP  621 bb.      |  Z     3.
15 4  DOQDO    Ca   1   2  14   1  -1  -1  -1 ..... 5 1edm  0 0.2 1.5  CA    1 ASPC  47 ...b.   |
16 1  DOQDO    Ca   1   2  14   1  -1  -1  -1 ..... 5 1edm  0 0.2 1.5  CA    1 ASPC  47 ...b.   |
17 1  DOOHDE   Ca   1  11   0   3   1  -1  -1 .. d..6 1gcy  0 0.2 1.6  CA  452 ASPA   1 b.....  |        3.
18 1  DOOODE   Ca   1   2   2  18   3  -1  -1 ......6 1i76  0 0.1 1.2  CA  997 ASPA 154 ......  |        3.
19 3  ODDODO   Ca   0   3   1   2   2  -1  -1 ......7 1wdc  0 0.2 2.0  CA  501 ASPC  19 ......  |      Z



over chelate loop

pchains 1-17 are similar, but within there are subgroups


p chain 1-4

relseq      meanfi             meanpsi       number

     0         -96( 39)         120( 11)           4         b

     1       -105(  4)         154(  7)           4         b


p chain 5-9

 relseq      meanfi             meanpsi       number

     0       -101( 36)         140( 10)           5         b

     1         -95( 13)         126(  6)           5         b


p chain 10-13

  relseq      meanfi             meanpsi       number

     0       -106( 41)          97( 12)           4         b

     1         -72(  5)         141(  8)           4         b


p chain 14 15
```

```
 relseq      meanfi             meanpsi       number
      0         -86(   8)        137(   1)           2         b
      1        -120(  1)        -166(  1)           2         b


all 1-17
 relseq      meanfi             meanpsi       number
      0         -95( 35)         124( 20)          16         b
      1         -97( 19)         148( 25)          17         b


no similarities in local conformation
```

```
Ca   OD 0
    (both donors belong to same aspartate residue)


   amino-acid sequences ****
    start of selected part of cngroup         |
    relseq                -10         -5          0          5          10          15
 1  1byf from   98 B:  Q I W S K Y N L L D D V G C G G A R R V I
 2  2msb from  196 A:  V T I V D N G L W N D I S C Q A S H T A V
 3  2msb from  196 B:  V T I V D N G L W N D I S C Q A S H T A V
 4  1egi from  738 A:  L K G D P T M S W N D I N C E H L N N W I
 5  1tn3 from  155  :  L S G A A N G K W F D K R C R D Q L P Y I
 6  4sgb from  110 E:  N T T P _ _ _ _ _ _ D I T S A _ _ _ _ _ A
 7  1e43 from  184 A:  W E V S S E N G N Y D Y L M Y A D V D Y D
 8  1e29 from   25 A:  Q F T N G Q K I F V D T C T Q C H L Q G K
 9  1wdc from    9 C:  D D L K D V F E L F D F W D G R D G A V D
10  1aru from   47  :  V R K I L R I V F H D A I G F S P A L T A
11  1pa2 from   33 A:  G A S L I R L H F H D C F V N G C D A S I
12  1e8u from  251 A:  C S K V T E T E E E D Y N S A V P T L M A
13  1e8u from  251 B:  C S K V T E T E E E D Y N S A V P T L M A
14  1d2v from   86 A:  F M Q W G Q L L D H D L D F T P E P _ _ _


   aa types - acidic, etc **
    relseq                -10         -5          0          5          10          15
 1  1byf from   98 B:  p h h s b h p h h a a h g h g g s b b h h
 2  2msb from  196 A:  h s h h a p g h h p a h s h p s s p s s h
 3  2msb from  196 B:  h s h h a p g h h p a h s h p s s p s s h
 4  1egi from  738 A:  h b g a h s h s h p a h p h a p h p p h h
 5  1tn3 from  155  :  h s g s s p g b h h a b b h b a p h h h h
 6  4sgb from  110 E:  p s s h _ _ _ _ _ _ a h s s s _ _ _ _ _ s
 7  1e43 from  184 A:  h a h s s a p g p h a h h h h s a h a h a
 8  1e29 from   25 A:  p h s p g p b h h h a s h s p h p h p g b
 9  1wdc from    9 C:  a a h b a h h a h h a h h a g b a g s h a
```

```
10  1aru from   47  :  h b b h h b h h h p a s h g h s h s h s s
11  1pa2 from   33 A:  g s s h h b h p h p a h h h p g h a s s h
12  1e8u from  251 A:  h s b h s a s a a a a h p s s h h s h h s
13  1e8u from  251 B:  h s b h s a s a a a a h p s s h h s h h s
14  1d2v from   86 A:  h h p h g p h h a p a h a h s h a h _ _ _


   conformation sequences **
    relseq            -10       -5         0         5        10        15
 1  1byf from   98 B:  b b b a a k g b b b b b b k k b b b b b b
 2  2msb from  196 A:  b b b b a k g b b b b b b k k b b b b b b
 3  2msb from  196 B:  b b b . b g g b b b b b b k k b b b b b b
 4  1egi from  738 A:  b b d b a a g b b b b b b k k b b b b b b
 5  1tn3 from  155  :  b b a a k g g b b b b b b a k b b b b b b
 6  4sgb from  110 E:  b a d b . . . . . . b b a b b . . . . . b
 7  1e43 from  184 A:  b b b b k b g . b k a b a b . b b b b k a
 8  1e29 from   25 A:  a a a a a a a a a a a a a k a k d b g g b
 9  1wdc from    9 C:  a a a a a a a a a a a a k k . b . g b b b
10  1aru from   47  :  a a a a a a a a a a a a k d b b a a a a a
11  1pa2 from   33 A:  a a a a a a a a a a a a k b g j d k g a a
12  1e8u from  251 A:  b b b b k b b a a a a a k b a b b b b b b
13  1e8u from  251 B:  b b b b k b b a a a a k k b a b b b b b b
14  1d2v from   86 A:  a a a a a a a a a k a a b b b b b b . . .


 **********  protein names (pdb header) *******
 1  1byf at  108 B   2.00     SUGAR BINDING PROTEIN  C-type lectin-like
 2  2msb at  206 A   1.7      LECTIN     C-type lectin-like
 3  2msb at  206 B   1.7      LECTIN       C-type lectin-like
 4  1egi at  748 A   2.30     SUGAR BINDING PROTEIN   C-type lectin-like
 5  1tn3 at  165      2.0      LECTIN       C-type lectin-like
 6  4sgb at  120 E   2.1      COMPLEX(SERINE PROTEINASE-INHIBITOR)Trypsin-like serine proteases
 7  1e43 at  194 A   1.7      HYDROLASE  TIM beta/alpha-barrel /alpha-Amylases, N-terminal domain
 8  1e29 at   35 A    1.21     ELECTRON TRANSPORT  Cytochrome c
 9  1wdc at   19 C    2.0      MUSCLE PROTEIN   EF Hand-like
```

```
10  1aru at   57     1.6      PEROXIDASE (DONOR:H2O2 OXIDOREDUCTAS  Heme-dependent peroxidases

11  1pa2 at   43 A   1.45     OXIDOREDUCTASE    Heme-dependent peroxidase

12  1e8u at  261 A   2.0      SIALIDASE         6-bladed beta-propeller

13  1e8u at  261 B   2.0      SIALIDASE

14  1d2v at   96 A   1.75     OXIDOREDUCTASE    Heme-dependent peroxidases


  **********  cngps  ************

  1 4  ENDOD     Ca   3  18   1   0  -1  -1  -1 ..... 7 1byf  0 0.1 2.0  CA  201 GLUB  86 .....   |    UU

  2 5  ENENOD   Ca   2   6  12   1   0  -1  -1 ......6 2msb  0 0.2 1.7  CA    2 GLUA 185 ......  |

  3 5  ENENOD   Ca   2   6  12   1   0  -1  -1 ......8 2msb  0 0.2 1.7  CA    2 GLUB 185 ......  |     UU

  4 2  NOD       Ca   1   0  -1  -1  -1  -1  -1 ...   4 1egi  0 0.2 2.3  CA  801 ASNA 747 ...     |  Z

  5 3  QEOD      Ca   7  15   0  -1  -1  -1  -1 ....  5 1tn3  2 0.2 2.0  CA  183 GLN  143 ....    |   Z

  6 1  ODOOY     Ca   0   1 122   0  -1  -1  -1 ..... 6 4sgb  0 0.1 2.1  CA    8 GLYE 120 .....   |    Z

  7 2  NODDO     Ca  92   0   6  35  -1  -1  -1 ....  6 1e43  1 0.2 1.7  CA  501 ASNA 102 .....   |    Z  3.

  8 1  OD         Ca   0  -1  -1  -1  -1  -1  -1 ..    6 1e29  0 0.3 1.2  CA  226 ASPA  35 ..      | ZZZZ

  9 1  ODDODO   Ca   0   3   1   2   2  -1  -1 ......7 1wdc  0 0.2 2.0  CA  501 ASPC  19 ......  |     Z

 10 1  ODODS     Ca   0  18   2   2  -1  -1  -1 ..... 7 1aru  0 0.2 1.6  CA  346 ASP   57 .....   |    ZZ 1.

 11 1  ODOODS   Ca   0   3   2   2   2  -1  -1 ......7 1pa2  0 0.1 1.5  CA  307 ASPA  43 ......  |     Z 1.

 12 1  ODOOO     Ca   0   3   2  30  -1  -1  -1 ..... 5 1e8u  0 0.1 2.0  CA 1002 ASPA 261 .....   |        3.

 13 1  ODOSOO   Ca   0   3   0   2  30  -1  -1 ......6 1e8u  0 0.2 2.0  CA 1003 ASPB 261 ......  |        3.

 14 1  ODOTODS  Ca   0 -99   0   2   2   2  -1 ......7 1d2v  0 0.2 1.8  CA  600 ASPA  96 ....... |        1.



note 2 and 3 have fairly different conformations over -10 to 10 - possibly

          peptide flips at relseq -5 and -6

      12 and 13 have fairly different conformations over -10 to 10, around relseq -5, -6

                       so say only 12 unique


group 1   pc 1-6    model:   2  2msb at  206 A   1.7


  relseq     meanfi          meanpsi       number

      0        -89( 15)         144(  7)           6         b

                 (without 6 agreement is much better)
```

```
                  local conformations fairly different (30-80 deg)


group 2  pc 7-14    model:  1pa2 at   43 A   1.45


  relseq      meanfi            meanpsi       number

     0        -68( 10)         -45(  9)         8         a

                   local conformations different, except 12 13 similar




group 1  pc 1-5 chi1,chi2  all within 25 deg range

           pc   6    chi1 different


group 2 pc 7-16   chi1   all within 17 deg

                        chi2   -46 to +32
```

```
Ca OD 2


  amino-acid sequences ****

   start of selected part of cngroup        |

   relseq              -10        -5         0          5        10        15

 1  1kap from  353 P:  A G N D I L Y G G L G A D Q L W G G A G A D

 2  1cdl from   52 C:  I N E V D A D G N G T I D F P E F L T M M A

 3  1kap from  278 P:  K L V F S V W D A G G N D T L D F S G F S Q

 4  1kap from  362 P:  L G A D Q L W G G A G A D T F V Y G D I A E

 5  1kap from  326 P:  V T V E N A I G G S G S D L L I G N D V A N

 6  1kap from  344 P:  D V A N V L K G G A G N D I L Y G G L G A D

 7  1dyk from 2864 A:  S Q T I S P K K A D I L D V V G I L Y V G G

 8  1dyk from 3043 A:  D A Q S P N S A S T S A D T N D P V F V G G

 9  1i8a from    4 A:   A K Y G T P V I D G E I D E I W N T T E E I

10  1d0b from   39 A:  T V S T P I K Q I F P D D A F A E T I K D N

11  1e43 from  171 A:  R I F K F R G E G K A W D W E V S S E N G N

12  1clc from  231  :  A L E I P E K N N S I P D F L D E L K Y E I

13  1tf4 from  494 B:  L D K G T F R Y W F T L D E G V D P A D I T

14  1ajj from   17  :   C I H S S W R C D G G P D C K D K S D E E N

15  2sic from  185 E:  Q R A S F S S V G P E L D V M A P G V S I Q

16  1kap from  442 P:  S L A I D F S G D A H A D F A I N L I G Q A

17  1ava from  136 A:  N P D T G A D F G A A P D I D H L N L R V Q

18  1djx from  641 A:  L P K V N K N K N S I V D P K V I V E I H G

19  1bag from  159  :  K R F L E R A L N D G A D G F R F D A A K H

20  1lpb from  180 B:  P C F Q G T P E L V R L D P S D A K F V D V

21  1qho from   38 A:  P T K S K W K M Y W G G D L E G V R Q K L P

22  1ga6 from  336 A:  Y G G Y G Y N A G T G W D Y P T G W G S L D

23  1sra from  215  :  H W Q F G Q L D Q H P I D G Y L S H T E L A

24  1i8a from    0 A:   _ V A T A K Y G T P V I D G E I D E I W N T

25  1pa2 from  217 A:  I T N L D L S T P D A F D N N Y F A N L Q S

26  1wdc from   13 C:  D V F E L F D F W D G R D G A V D A F K L G

27  1aru from  197  :  F R S P L D S T P Q V F D T Q F Y I E T L L
```

```
28  1pa2 from   38 A:  R L H F H D C F V N G C D A S I L L D D T G

29  1cru from  261 B:  G N Y G W P N V A G Y K D D S G Y A Y A N Y

30  1aru from   65  :  L T A A G Q F G G G G A D G S I I A H S N I

31  1qho from   67 A:  T T I W L S P V L D N L D T L A G T D N T G

32  1i76 from  161 A:  A H A F Q P G Q G I G G D A H F D A E E T W


   aa types - acidic, etc **

    relseq                -10         -5          0          5         10         15

 1   1kap from  353 P:  s g p a h h h g g h g s a p h h g g s g s a

 2   1cdl from   52 C:  h p a h a s a g p g s h a h h a h h s h h s

 3   1kap from  278 P:  b h h h s h h a s g g p a s h a h s g h s p

 4   1kap from  362 P:  h g s a p h h g g s g s a s h h h g a h s a

 5   1kap from  326 P:  h s h a p s h g g s g s a h h h g p a h s p

 6   1kap from  344 P:  a h s p h h b g g s g p a h h h g g h g s a

 7   1dyk from 2864 A:  s p s h s h b b s a h h a h h g h h h h g g

 8   1dyk from 3043 A:  a s p s h p s s s s s s a s p a h h h h g g

 9   1i8a from    4 A:  s b h g s h h h a g a h a a h h p s s a a h

10  1d0b from   39 A:  s h s s h h b p h h h a a s h s a s h b a p

11  1e43 from  171 A:  b h h b h b g a g b s h a h a h s s a p g p

12  1clc from  231  :  s h a h h a b p p s h h a h h a a h b h a h

13  1tf4 from  494 B:  h a b g s h b h h h s h a a g h a h s a h s

14  1ajj from   17  :  h h p s s h b h a g g h a h b a b s a a a p

15  2sic from  185 E:  p b s s h s s h g h a h a h h s h g h s h p

16  1kap from  442 P:  s h s h a h s g a s p s a h s h p h h g p s

17  1ava from  136 A:  p h a s g s a h g s s h a h a p h p h b h p

18  1djx from  641 A:  h h b h p b p b p s h h a h b h h h a h p g

19  1bag from  159  :  b b h h a b s h p a g s a g h b h a s s b p

20  1lpb from  180 B:  h h h p g s h a h h b h a h s a s b h h a h

21  1qho from   38 A:  h s b s b h b h h h g g a h a g h b p b h h

22  1ga6 from  336 A:  h g g h g h p s g s g h a h h s g h g s h a

23  1sra from  215  :  p h p h g p h a p p h h a g h h s p s a h s

24  1i8a from    0 A:  _ h s s s b h g s h h h a g a h a a h h p s

25  1pa2 from  217 A:  h s p h a h s s h a s h a p p h h s p h p s
```

```
26  1wdc from   13 C: a h h a h h a h h a g b a g s h a s h b h g
27  1aru from  197  : h b s h h a s s h p h h a s p h h h a s h h
28  1pa2 from   38 A: b h p h p a h h h p g h a s s h h h a a s g
29  1cru from  261 B: g p h g h h p h s g h b a a s g h s h s p h
30  1aru from   65  : h s s s g p h g g g g s a g s h h s p s p h
31  1qho from   67 A: s s h h h s h h h a p h a s h s g s a p s g
32  1i76 from  161 A: s p s h p h g p g h g g a s p h a s a a s h


   conformation sequences **
    relseq          -10        -5         0         5        10        15
 1  1kap from  353 P: b j b b b b b a j b j b b b b b a j b j b b
 2  1cdl from   52 C: a a a a b a k g k g b b b a a a a a a a a k
 3  1kap from  278 P: b b b . b b b b k . j b b b b b d a k b k b
 4  1kap from  362 P: b j b b b b b a j b j b b b b b b k b a a a
 5  1kap from  326 P: b b b a b b b b . k b b b b b b b b k b b b
 6  1kap from  344 P: k b b b b b b a j b j b b b b b a j b j b b
 7  1dyk from 2864 A: b b . b b b . k b k b b b b b d b b b b g .
 8  1dyk from 3043 A: b b b b k d k a b k b b d b . b b b b b g .
 9  1i8a from    4 A: b b b . b b b b . g b b b a a a k k b b b b
10  1d0b from   39 A: k b b b b a k a a b a b a a a a a a a a a a
11  1e43 from  171 A: b b b b b b . b g b . b b b b b b k b g . b
12  1clc from  231  : b b d b a k k k b k b b a a a a a a a a a a
13  1tf4 from  494 B: b k d b b b b b b b b b b b g b b a k k b b
14  1ajj from   17  : b b b a k k b g k g b b a b a k g a k k b g
15  2sic from  185 E: b b b a a b b b . k k b a b b b b k b g b b
16  1kap from  442 P: . b b b b a a g k g b b a b b b b b b j b b
17  1ava from  136 A: b b b k . b b b a k b b b b b k k b a a a a
18  1djx from  641 A: b b b d g . . . a b b b d b b b b b b b b .
19  1bag from  159  : a k a a a a a a a k . b k . b b b b g k k k
20  1lpb from  180 B: b g . b g b b a a k b b . a k k b a b b b b
21  1qho from   38 A: k k g k b k k b b k g b b a a a a a a k a a
22  1ga6 from  336 A: b g g b j b b b . b j b b a a k j k j b b b
23  1sra from  215  : a a a a a a a b k d b b . g b b b a a k k a
```

```
24  1i8a from    0 A:  . b b b b b b . b b b b . g b b b a a a k k

25  1pa2 from  217 A:  b b b a b k a b a k b b b k a a a a a a k k

26  1wdc from   13 C:  a a a a a a a a k k . b . g b b b k k k a a

27  1aru from  197  :  b g b b k b k d k k b b b k a a a a a k k b

28  1pa2 from   38 A:  a a a a a a a k b g j d k g a a k b b b b a

29  1cru from  261 B:  b b k g b k a b b j b b k g b g b b b b b a

30  1aru from   65  :  a a a k g b b k . k j d k g a a a a d a a a

31  1qho from   67 A:  k b b b b b b k b b b b k b b d j b g b b d

32  1i76 from  161 A:  . b b b b b . b . a b g b b b b b k k b b b


  **********  protein names (pdb header) *******

 1  1kap at  363 P   1.64     ZINC METALLOPROTEASE      beta-Roll  Metalloprotease, C-terminal domain

 2  1cdl at   62 C   2.2      CALCIUM-BINDING PROTEIN    EF Hand-like Calmodulin-like

 3  1kap at  288 P   1.64     ZINC METALLOPROTEASE      beta-Roll

 4  1kap at  372 P   1.64     ZINC METALLOPROTEASE      beta-Roll

 5  1kap at  336 P   1.64     ZINC METALLOPROTEASE      beta-Roll

 6  1kap at  354 P   1.64     ZINC METALLOPROTEASE      beta-Roll

 7  1dyk at 2874 A   2.0      METAL BINDING PROTEIN     Concanavalin A-like lectins/glucanases

 8  1dyk at 3053 A   2.0      METAL BINDING PROTEIN             Laminin G-like module

 9  1i8a at   14 A   1.90     HYDROLASE                   Immunoglobulin-like beta-sandwich

                                                                     Family 9 carbohydrate-binding module, CBD9

10  1d0b at   49 A   1.86     CELL ADHESION    Leucine-rich repeat, LRR (right-handed beta-alpha sup

11  1e43 at  181 A   1.7      HYDROLASE        TIM beta/alpha-barrel

12  1clc at  241     1.9      GLYCOSYL HYDROLASE          alpha/alpha toroid  Cellulases catalytic domain


13  1tf4 at  504 B   1.9      GLYCOSYL HYDROLASE Common fold of diphtheria toxin/transcription factors/cytochrome f

                                                                 Cellulose-binding domain family III

14  1ajj at   27     1.7      RECEPTOR

15  2sic at  195 E   1.8      COMPLEX (PROTEINASE/INHIBITOR)

16  1kap at  452 P   1.64     ZINC METALLOPROTEASE

17  1ava at  146 A   1.9      HYDROLASE INHIBITION


18  1djx at  651 A   2.30     LIPID DEGRADATION
```

```
19  1bag  at   169       2.5        ALPHA-AMYLASE

20  1lpb  at   190  B    2.46       HYDROLASE(CARBOXYLIC  ESTERASE)

21  1qho  at    48  A    1.70       HYDROLASE

22  1ga6  at   346  A    1.00       HYDROLASE


23  1sra  at   225       2.0        CALCIUM-BINDING  PROTEIN

24  1i8a  at    10  A    1.90       HYDROLASE

25  1pa2  at   227  A    1.45       OXIDOREDUCTASE

26  1wdc  at    23  C    2.0        MUSCLE  PROTEIN

27  1aru  at   207       1.6        PEROXIDASE  (DONOR:H2O2  OXIDOREDUCTAS


28  1pa2  at    48  A    1.45       OXIDOREDUCTASE

29  1cru  at   271  B    1.50       OXIDOREDUCTASE


30  1aru  at    75       1.6        PEROXIDASE  (DONOR:H2O2  OXIDOREDUCTAS

31  1qho  at    77  A    1.70       HYDROLASE

32  1i76  at   171  A    1.20       HYDROLASE


  **********   cngps   ************

  1  2   OODOD      Ca    2    2   18    7   -1   -1   -1  ..... 6  1kap   0  0.1  1.6   CA   619  GLYP  361  .....    |      Z   3.

  2  4   DDNODE    Ca    2    2    2    2    3   -1   -1  ......6  1cdl   0  0.6  2.2   CA     2  ASPC   56  .....b   |

  3  1   ODOE       Ca    2   37    2   -1   -1   -1   -1  ....   6  1kap   0  0.1  1.6   CA   615  GLYP  288  ...b     |    ZZ   3.

  4  2   OODD       Ca    2    2   26   -1   -1   -1   -1  ....   6  1kap   1  0.2  1.6   CA   617  GLYP  370  ....     |    ZZ   3.

  5  2   OODOOD    Ca    2    2   13    2    3   -1   -1  ......6  1kap   0  0.3  1.6   CA   616  GLYP  334  .....b   |         3.

  6  2   OODOOD    Ca    2    2   13    2    3   -1   -1  ......6  1kap   0  0.3  1.6   CA   620  GLYP  352  .....b   |         3.

  7  3   DOOD       Ca   17   49    2   -1   -1   -1   -1  ....   5  1dyk   0  0.2  2.0   CA  4001  ASPA2808  ....     |    Z

  8  3   DOOD       Ca   17   54    2   -1   -1   -1   -1  ....   4  1dyk   0  0.2  2.0   CA  4002  ASPA2982  ....     |

  9  3   ODODE      Ca    2    2    2  114   -1   -1   -1  ..... 6  1i8a   0  0.2  1.9   CA   191  VALA   10  ....b    |      Z   3.

 10  1   OD          Ca    2   -1   -1   -1   -1   -1   -1  ..      7  1d0b   0  0.2  1.9   CA   201  PROA   49  ..        |  ZZZZZ

 11  2   DODDD      Ca   22    2   19    2   -1   -1   -1  ..... 6  1e43   0  0.1  1.7   CA   502  ASPA  159  b....    |      Z   3.

 12  3   ONODD      Ca    3    2    2    3   -1   -1   -1  ..... 6  1clc   0  0.1  1.9   CA   591  GLU   236  .....    |      Z   3.

 13  1   ODOND      Ca    2   65    3    1   -1   -1   -1  ..... 6  1tf4   0  0.1  1.9   CA  3004  THRB  504  .b...    |      Z   3.

 14  3   ODODDE    Ca    3    2    2    6    1   -1   -1  ......6  1ajj   0  0.1  1.7   CA    73  TRP    22  ......   |
```

```
15  3   OOOD       Ca    5   21    2   -1   -1   -1   -1 ....   6  2sic   1  0.6  1.8   CA   502  GLYE  169  ....     |    ZZ   3.
16  4   DSDOD      Ca    2    2    2    2   -1   -1   -1 ... .  6  1kap   0  0.3  1.6   CA   621  ASPP  446  ....b    |     Z   3.
17  4   DDOOD      Ca   15    1    3    2   -1   -1   -1 ..... 6  1ava   1  0.2  1.9   CA   502  ASPA  127  .....    |     Z   3.
18  1   ODN        Ca    2   24   -1   -1   -1   -1   -1 ...    5  1djx   1  0.4  2.3   CA     3  ILEA  651  .b.       |   ZZ    3.
19  1   OD         Ca    2   -1   -1   -1   -1   -1   -1 ..     5  1bag   0  0.1  2.5   CA     7  GLY   169  .b        | ZZZ    3.
20  2   OODD       Ca    3    2    3   -1   -1   -1   -1 ....   6  1lpb   0  0.2  2.5   CA   450  GLUB  187  ...b     |    ZZ   3.
21  5   DONNOD     Ca    2    3    1   21    2   -1   -1 ......7  1qho   0  0.2  1.7   CA   698  ASPA   21  ......   |      Z  3.
22  4   DOOOD      Ca    1   15    2    2   -1   -1   -1 ..... 6  1ga6   0  0.1  1.0   CA   374  ASPA  328  .....    |     Z   3.
23  2   DODOE      Ca    3    2    2    5   -1   -1   -1 ..... 6  1sra   0  0.2  2.0   CA   301  ASP   222  ....b    |     Z
24  1   ODODE      Ca    2    2    2  114   -1   -1   -1 ..... 6  1i8a   0  0.2  1.9   CA   191  VALA   10  ....b    |     Z   3.
25  6   OTDOTOD    Ca    0   51    3    0    3    2   -1 ......7  1pa2   0  0.1  1.5   CA   308  THRA  170  ....... |         1.
26  4   ODDODO     Ca    0    3    1    2    2   -1   -1 ......7  1wdc   0  0.2  2.0   CA   501  ASPC   19  ......   |      Z
27  6   OSDOTOD    Ca    0   17    2    0    3    2   -1 ......7  1aru   0  0.2  1.6   CA   347  SER   185  ..b.... |         1.
28  4   ODOODS     Ca    0    3    2    2    2   -1   -1 ......7  1pa2   0  0.1  1.5   CA   307  ASPA   43  ......   |      Z  1.
29  2   OODE       Ca    2    2   36   -1   -1   -1   -1 ....   6  1cru   0  0.1  1.5   CA   902  ALAB  269  ...b     |    ZZ   1.
30  3   ODODS      Ca    0   18    2    2   -1   -1   -1 ..... 7  1aru   0  0.2  1.6   CA   346  ASP    57  .....    |     ZZ  1.
31  2   DODEE      Ca    1    2   22    1   -1   -1   -1 ..... 6  1qho   1  0.2  1.7   CA   696  ASPA   76  ....b    |     Z   3.
32  3   OOOD       Ca   32    2    2   -1   -1   -1   -1 ....   6  1i76   0  0.1  1.2   CA   996  ASPA  137  ....     |    ZZ   3.
```

seems to be a gradation in conformations ?? over fi 1-2, psi 1 1

pchain 1-7

```
   relseq      meanfi          meanpsi       number
      1          -90(   4)       106(  10)          7         b
      2         -101(   8)       149(  20)          7         b
```

pc 8-9

```
   relseq      meanfi          meanpsi       number
      1          -89(   5)        98(   8)          2         b
      2          -66(   7)       116(  54)          2         b
```

```
pchain 10-12

  relseq      meanfi           meanpsi       number

      1         -57(  7)        132( 11)          3         b

      2         -58(  9)        -93( 99)          3         a


pchain 13-17

  relseq      meanfi           meanpsi       number

      1         -64(  3)        130(  5)          5         b

      2         -99(  9)        253( 93)          5         .


pchain 18-22

  relseq      meanfi           meanpsi       number

      1         -62(  4)        149(  8)          5         b

      2       -142( 16)        115( 79)          5         b


pchain 23-27

relseq      meanfi           meanpsi       number

      1         -78(  7)        105(  7)          5         b

      2       -149(  5)          77( 93)          5         d


pchain 28-29

  relseq      meanfi           meanpsi       number

      1         -83(  3)          80(  0)          2         d

      2       -127(  2)           5( 11)          2         k


all of these plus 30,31  i.e. pchains 1-31

  relseq      meanfi           meanpsi       number

      1         -73( 16)        116( 23)         31         b

      2       -110( 31)        128(103)         31         b



pchain 32 is entirely different
```

```
Ca OE 2   compared with Ca OD 2
   1-5 are caod2

  amino-acid sequences ****
   start of selected part of cngroup         |
   relseq               -10        -5         0         5        10        15
 1  1kap from  353 P:  A G N D I L Y G G L G A D Q L W G G A G A D
 2  1e43 from  171 A:  R I F K F R G E G K A W D W E V S S E N G N
 3  1ga6 from  336 A:  Y G G Y G Y N A G T G W D Y P T G W G S L D
 4  1pa2 from   38 A:  R L H F H D C F V N G C D A S I L L D D T G
 5  1i76 from  161 A:  A H A F Q P G Q G I G G D A H F D A E E T W

 6  2pvb from   47 A:  F Y V I D Q D K S G F I E E D E L K L F L Q
 7  1dx5 from  414 J:  I L D D G F I C T D I D E C E N G G F C S G
 8  1fjs from   65 A:  K V R V G D R N T E Q E E G G E A V H E V E
 9  1avw from   65 A:  V R _ _ _ E H N I D V L E G N E Q F I N A A
10  1kap from  317 P:  K G N V S I A A G V T V E N A I G G S G S D
11  1cvr from   95 A:  V Y G Q I V G N D H Y N E V F I G R F S C E
12  1c1y from  113 B:  D W N T D A A S L I G E E L Q V D F _ _ _ _
13  1dyo from    4 A:  _ _ D A G Y Y Y H D T F E G S V G Q W T A R

  aa types - acidic, etc **
   relseq               -10        -5         0         5        10        15
 1  1kap from  353 P:  s g p a h h h g g h g s a p h h g g s g s a
 2  1e43 from  171 A:  b h h b h b g a g b s h a h a h s s a p g p
 3  1ga6 from  336 A:  h g g h g h p s g s g h a h h s g h g s h a
 4  1pa2 from   38 A:  b h p h p a h h h p g h a s s h h h a a s g
 5  1i76 from  161 A:  s p s h p h g p g h g g a s p h a s a a s h

 6  2pvb from   47 A:  h h h h a p a b s g h h a a a a h b h h h p
 7  1dx5 from  414 J:  h h a a g h h h s a h a a h a p g g h h s g
 8  1fjs from   65 A:  b h b h g a b p s a p a a g g a s h p a h a
```

```
  9   1avw  from    65  A:   h  b  _  _  _  a  p  p  h  a  h  h  a  g  p  a  p  h  h  p  s  s

 10   1kap  from   317  P:   b  g  p  h  s  h  s  s  g  h  s  h  a  p  s  h  g  g  s  g  s  a

 11   1cvr  from    95  A:   h  h  g  p  h  h  g  p  a  p  h  p  a  h  h  h  g  b  h  s  h  a

 12   1c1y  from   113  B:   a  h  p  s  a  s  s  s  h  h  g  a  a  h  p  h  a  h  _  _  _  _

 13   1dyo  from     4  A:   _  _  a  s  g  h  h  h  p  a  s  h  a  g  s  h  g  p  h  s  s  b


    conformation  sequences  **

     relseq                 -10           -5            0            5           10           15

  1   1kap  from   353  P:   b  j  b  b  b  b  b  a  j  b  j  b  b  b  b  b  a  j  b  j  b  b

  2   1e43  from   171  A:   b  b  b  b  b  b  .  b  g  b  .  b  b  b  b  b  b  k  b  g  .  b

  3   1ga6  from   336  A:   b  g  g  b  j  b  b  b  .  b  j  b  b  a  a  k  j  k  j  b  b  b

  4   1pa2  from    38  A:   a  a  a  a  a  a  a  k  b  g  j  d  k  g  a  a  k  b  b  b  b  a

  5   1i76  from   161  A:   .  b  b  b  b  b  .  b  .  a  b  g  b  b  b  b  b  k  k  b  b  b


  6   2pvb  from    47  A:   a  a  a  a  d  a  k  g  k  g  b  b  b  a  a  a  a  a  k  a  a  a

  7   1dx5  from   414  J:   b  b  b  b  g  g  b  b  b  b  b  b  a  a  a  a  j  g  a  b  a  .

  8   1fjs  from    65  A:   b  b  b  b  g  b  k  b  a  d  .  b  a  g  k  b  b  b  b  b  b  a

  9   1avw  from    65  A:   b  b  .  .  .  b  a  b  k  k  b  b  k  g  k  b  b  b  b  b  b  a

 10   1kap  from   317  P:   b  g  k  b  b  b  b  b  g  b  b  b  a  b  b  b  b  .  k  b  b  b

 11   1cvr  from    95  A:   a  a  k  b  a  b  g  g  b  k  b  b  a  b  b  b  b  b  b  b  b  k

 12   1c1y  from   113  B:   b  k  k  b  b  a  a  k  k  b  g  b  b  b  b  b  b  b  .  .  .  .

 13   1dyo  from     4  A:   .  .  .  k  g  b  a  b  b  b  a  g  d  g  .  k  j  k  b  b  b  b


  **********   protein  names  (pdb  header)  *******

  1   1kap  at   363  P    1.64       ZINC  METALLOPROTEASE

  2   1e43  at   181  A    1.7        HYDROLASE

  3   1ga6  at   346  A    1.00       HYDROLASE

  4   1pa2  at    48  A    1.45       OXIDOREDUCTASE

  5   1i76  at   171  A    1.20       HYDROLASE


  6   2pvb  at    57  A    0.91       METAL  BINDING  PROTEIN

  7   1dx5  at   424  J    2.3        SERINE  PROTEINASE

  8   1fjs  at    75  A    1.92       BLOOD  CLOTTING
```

```
  9  1avw  at    75  A    1.75      COMPLEX  (PROTEINASE/INHIBITOR)
 10  1kap  at   327  P    1.64      ZINC  METALLOPROTEASE
 11  1cvr  at   105  A    2.00      HYDROLASE
 12  1c1y  at   123  B    1.90      SIGNALING  PROTEIN
 13  1dyo  at    14  A    2.1       CARBOHYDRATE-BINDING  MODULE

  **********   cngps   ************
  1 2   OODOD      Ca    2    2   18    7   -1   -1   -1 ..... 6  1kap   0  0.1  1.6   CA   619  GLYP  361  .....    |     Z   3.
  2 2   DODDD      Ca   22    2   19    2   -1   -1   -1 ..... 6  1e43   0  0.1  1.7   CA   502  ASPA  159  b....    |     Z   3.
  3 4   DOOOD      Ca    1   15    2    2   -1   -1   -1 ..... 6  1ga6   0  0.1  1.0   CA   374  ASPA  328  .....    |     Z   3.
  4 4   ODOODS    Ca    0    3    2    2    2   -1   -1 ......7  1pa2   0  0.1  1.5   CA   307  ASPA   43  ......   |      Z  1.
  5 3   OOOD       Ca   32    2    2   -1   -1   -1   -1 ....   6  1i76   0  0.1  1.2   CA   996  ASPA  137  ....     |    ZZ   3.

  6 4   DDSOEE    Ca    2    2    2    2    3   -1   -1 ......6  2pvb   0  0.1  0.9   CA   110  ASPA   51  .....b   |
  7 2   DOENOO    Ca    1    2   13    1    3   -1   -1 ......7  1dx5   0  0.2  2.3   CA  1001  ASPJ  423  b.....   |      Z  3.
  8 3   DOOEE      Ca    2    3    2    3   -1   -1   -1 ..... 6  1fjs   0  0.3  1.9   CA   507  ASPA   70  .....    |     Z   3.
  9 3   EOOEE      Ca    2    3    2    3   -1   -1   -1 ..... 6  1avw   0  0.1  1.8   CA   700  GLUA   70  .....    |     Z   3.
 10 3   ODOE       Ca    2   37    2   -1   -1   -1   -1 ....   6  1kap   0  0.1  1.6   CA   615  GLYP  288  ...b     |    ZZ   3.
 11 3   ODOE       Ca    3    2    2   -1   -1   -1   -1 ....   6  1cvr   0  0.2  2.0   CA   477  VALA  100  ...b     |    ZZ   3.
 12 1   OE          Ca    2   -1   -1   -1   -1   -1   -1 ..     4  1c1y   0  0.2  1.9   CA   173  GLYB  123  .b        | ZZ      2.
 13 1   OEOOD      Ca    2   23    3  107   -1   -1   -1 ..... 5  1dyo   0  0.4  2.1   CA   300  THRA   14  ....b    |

 over  chelate  loop
 pchains  1,6,7,8,9,10  are  pretty   alike,  for  6-10
  relseq      meanfi             meanpsi        number
       1         -90(  11)         112(  15)           5         b
       2         -86(  15)         300(  72)           5         a

 pchains  11  12  like  each  other  and  somewhat  like  2
  relseq      meanfi             meanpsi        number
       1         -68(  17)         152(   4)           2         b
       2       -106(   7)           49(108)           2         d
```

```
pchains 5 13 alike

 relseq       meanfi             meanpsi        number

      1           79(   0)            0( 24)            2          g

      2          -76(   4)          114( 52)            2          b


7, 8 are similar (rms 27), otherwise local conformations fairly different
```

```
Ca OE 5

    p chains 1 2 3 from 2225  EF hand type set, 4 5 6 different

     (remainder of 2225 set omitted)

   amino-acid sequences ****

    start of selected part of cngroup         |

    relseq               -10         -5          0          5         10         15

  1  1acc from  173  :  _ _ V P D R D N D G I P D S L E V E G Y T V D V K

  2  1g4y from   16 R:  F S L F D K D G D G T I T T K E L G T V M R S L G

  3  1g4y from   52 R:  I N E V D A D G N G T I D F P E F L T M M A R K M

  4  1sbw from   65 A:  V R L _ G E D N I N V V E G N E Q F I S A S K S I

  5  2btc from   65 E:  V R L _ G E D N I N V V E G N E Q F I S A S K S I

  6  1alv from  102 B:  L F A Q L A G D D M E V S A T E L M N I L N K V V


   aa types - acidic, etc **

    relseq               -10         -5          0          5         10         15

  1  1acc from  173  :  _ _ h h a b a p a g h h a s h a h a g h s h a h b

  2  1g4y from   16 R:  h s h h a b a g a g s h s s b a h g s h h b s h g

  3  1g4y from   52 R:  h p a h a s a g p g s h a h h a h h s h h s b b h

  4  1sbw from   65 A:  h b h _ g a a p h p h h a g p a p h h s s s b s h

  5  2btc from   65 E:  h b h _ g a a p h p h h a g p a p h h s s s b s h

  6  1alv from  102 B:  h h s p h s g a a h a h s s s a h h p h h p b h h


   conformation sequences **

    relseq               -10         -5          0          5         10         15

  1  1acc from  173  :  . . . b d a k g k g b b a a a a a a . b b b b b b

  2  1g4y from   16 R:  a a k a b a k g k g b b b a a a a a a a a a k k g

  3  1g4y from   52 R:  a a a d b a k g k g b b b a a a k a a a a a k d a

  4  1sbw from   65 A:  b b . . . b . b k k b b k . g b b b b b b a b b b

  5  2btc from   65 E:  b b . . . b a b k k b b k g k b b b b b b a b b b

  6  1alv from  102 B:  a a a a a a j a k g b b b a a a a a a a a a a a a


  **********  protein names (pdb header) *******
```

```
 1   1acc  at   183       2.1        TOXIN
 2   1g4y  at    26  R    1.6        SIGNALING  PROTEIN
 3   1g4y  at    62  R    1.6        SIGNALING  PROTEIN
 4   1sbw  at    75  A    1.80       HYDROLASE/HYDROLASE  INHIBITOR
 5   2btc  at    75  E    1.50       HYDROLASE/HYDROLASE  INHIBITOR
 6   1alv  at   112  B    1.9        CALCIUM  BINDING


 **********   cngps   ************
 1  4   DDDOE      Ca    2    2    2    5   -1   -1   -1  .....  6  1acc   1  0.2  2.1   CA   800  ASP   177  .....    |      Z
 2  4   DDDOE      Ca    2    2    2    5   -1   -1   -1  .....  6  1g4y   0  0.2  1.6   CA  1001  ASPR   20  ....b    |      Z
 3  4   DDNOE      Ca    2    2    2    5   -1   -1   -1  .....  5  1g4y   0  0.4  1.6   CA  1002  ASPR   56  ....b    |
 4  3   EOOE       Ca    2    3    5   -1   -1   -1   -1  ....   6  1sbw   0  0.2  1.8   CA   801  GLUA   70  ....     |    ZZ   3.
 5  3   EOOE       Ca    2    3    5   -1   -1   -1   -1  ....   6  2btc   0  0.2  1.5   CA   700  GLUE   70  ....     |    ZZ   3.
 6  3   ODOE       Ca    3    2    5   -1   -1   -1   -1  ....   6  1alv   0  0.1  1.9   CA     5  ALAB  107  ...b     |    ZZ   3.
```

p chains 2 3 6 are similar at chelate loop - i.e. 6 in 1alv is like the end of 2225 set, but 4 and 5 are different from this and from each other

over loop 0 4, 1-5 p chains 2 3 6

```
 relseq       meanfi            meanpsi        number
      0        -139(   7)         155(   7)           3         b
      1        -101(   6)         122(   6)           3         b
      2          -97(  12)        170(   8)           3         b
      3          -54(  14)        -40(  10)           3         a
      4          -61(   7)         -42(   2)          3         a
      5          -68(   4)         -36(   6)          3         a
```

local conformation of 6 is similar to 2 over relseq -2 to 12

```
Ca    OS 0  and  OT 0


    amino-acid  sequences  ****

      start  of  selected  part  of  cngroup          |

      relseq                  -10          -5           0           5          10          15

  1   1kit  from   303   :   T  V  L  V  S  Y  A  R  W  P  T  D  A  A  Q  N  G  D  R  I  K

  2   1dl2  from   515  A:   E  F  D  L  T  K  V  V  F  N  T  E  A  H  P  F  P  V  L  D  E

  3   1pa2  from   160  A:   N  D  L  V  A  L  S  G  A  H  T  F  G  R  A  R  C  G  V  F  N

  4   1aru  from   175   :   D  E  V  V  D  L  L  A  A  H  S  L  A  S  Q  E  G  L  N  S  A

  5   1d2v  from   158  C:   I  T  I  R  N  Q  I  N  A  L  T  S  F  V  D  A  S  M  V  Y  G

  6   1jg8  from     4  C:   _  _  I  D  L  R  S  D  T  V  T  K  P  T  E  E  M  R  K  A  _

  7   1scj  from   164  A:   T  V  G  Y  P  A  K  Y  P  S  T  I  A  V  G  A  V  N  S  S  N

  8   1e8u  from   254  B:   V  T  E  T  E  E  E  D  Y  N  S  A  V  P  T  L  M  A  H  G  R

  9   1nbc  from    34  B:   L  S  K  L  T  L  R  Y  Y  Y  T  V  D  G  Q  K  D  Q  T  F  W

 10   1iod  from    31  A:   T  K  Q  V  N  G  G  H  L  V  S  I  E  S  S  G  E  A  D  F  V

 11   1pa2  from   214  A:   A  S  T  I  T  N  L  D  L  S  T  P  D  A  F  D  N  N  Y  F  A

 12   1aru  from   194   :   S  A  I  F  R  S  P  L  D  S  T  P  Q  V  F  D  T  Q  F  Y  I

 13   1edm  from    43  B:   _  _  _  _  D  G  D  Q  C  E  S  N  P  C  L  N  G  G  S  C  K


    aa  types  -  acidic,  etc  **

      relseq                  -10          -5           0           5          10          15

  1   1kit  from   303   :   s  h  h  h  s  h  s  b  h  h  s  a  s  s  p  p  g  a  b  h  b

  2   1dl2  from   515  A:   a  h  a  h  s  b  h  h  h  p  s  a  s  p  h  h  h  h  h  a  a

  3   1pa2  from   160  A:   p  a  h  h  s  h  s  g  s  p  s  h  g  b  s  b  h  g  h  h  p

  4   1aru  from   175   :   a  a  h  h  a  h  h  s  s  p  s  h  s  s  p  a  g  h  p  s  s

  5   1d2v  from   158  C:   h  s  h  b  p  p  h  p  s  h  s  s  h  h  a  s  s  h  h  h  g

  6   1jg8  from     4  C:   _  _  h  a  h  b  s  a  s  h  s  b  h  s  a  a  h  b  b  s  _

  7   1scj  from   164  A:   s  h  g  h  h  s  b  h  h  s  s  h  s  h  g  s  h  p  s  s  p

  8   1e8u  from   254  B:   h  s  a  s  a  a  a  a  h  p  s  s  h  h  s  h  h  s  p  g  b

  9   1nbc  from    34  B:   h  s  b  h  s  h  b  h  h  h  s  h  a  g  p  b  a  p  s  h  h

 10   1iod  from    31  A:   s  b  p  h  p  g  g  p  h  h  s  h  a  s  s  g  a  s  a  h  h

 11   1pa2  from   214  A:   s  s  s  h  s  p  h  a  h  s  s  h  a  s  h  a  p  p  h  h  s
```

```
12  1aru from  194  :  s s h h b s h h a s s h p h h a s p h h h
13  1edm from   43 B:  _ _ _ _ a g a p h a s p h h h p g g s h b


   conformation sequences **
    relseq                -10        -5         0         5        10        15
 1  1kit from  303  :  b b b b b b b b b b k k b . a b g g k b b
 2  1dl2 from  515 A:  b b b k k k b b b b k k g b b b b b b b a
 3  1pa2 from  160 A:  a a a a a a a k k k k b j b b b a a k k k
 4  1aru from  175  :  a a a a a a k a a k k d b b b a k a b k a
 5  1d2v from  158 C:  a b b b b b b b k b b a b b k g k k a k .
 6  1jg8 from    4 C:  . . b b k a . k a k b b b b a a a a a a .
 7  1scj from  164 A:  k b b b k a a b a a b b b b b b b b k k g
 8  1e8u from  254 B:  b k b b a a a a k k b a b b b b b b b b b
 9  1nbc from   34 B:  a k k b b b b b b . b a g . b b b b b b b
10  1iod from   31 A:  k k a b b g b b b b b b k . a a a a a a a
11  1pa2 from  214 A:  k k b b b b a b k a b a k b b b k a a a a
12  1aru from  194  :  k a d b g b b k b k d k k b b b k a a a a
13  1edm from   43 B:  . . . . b b b k k b g g a b b g g b b b b


 **********  protein names (pdb header) *******
 1  1kit at  313      2.3       HYDROLASE     Concanavalin A-like lectins/glucanases
 2  1dl2 at  525 A   1.54      HYDROLASE      alpha/alpha toroid
 3  1pa2 at  170 A   1.45      OXIDOREDUCTASE     Heme-dependent peroxidases
 4  1aru at  185      1.6       PEROXIDASE (DONOR:H2O2 OXIDOREDUCTAS Heme-dependent peroxidases
 5  1d2v at  168 C   1.75      OXIDOREDUCTASE    Heme-dependent peroxidases
 6  1jg8 at   14 C   1.80      LYASE  PLP-dependent transferases
 7  1scj at  174 A   2.0       HYDROLASE   Subtilisin-like
 8  1e8u at  264 B   2.0       SIALIDASE   6-bladed beta-propeller
 9  1nbc at   44 B   1.8       CELLULOSE DEGRADATION     Common fold of diphtheria toxin/transcripti
10  1iod at   41 A   2.30      HYDROLASE/HYDROLASE INHIBITOR     C-type lectin-like
11  1pa2 at  224 A   1.45      OXIDOREDUCTASE
12  1aru at  204      1.6       PEROXIDASE (DONOR:H2O2 OXIDOREDUCTAS
13  1edm at   53 B   1.5       COAGULATION FACTOR
```

```
 **********  cngps  ************
 1 5  OONDOT    Ca    3    0   33   24    0   -1   -1 ......6 1kit   0 0.2 2.3   CA   802 ALA   253 ...b..   |         3.
 2 1  OT         Ca    0   -1   -1   -1   -1   -1   -1 ..     6 1dl2   1 0.1 1.5   CA   901 THRA 525 ..        | ZZZZ  3.
 3 1  OTDOTOD   Ca    0   51    3    0    3    2   -1 ......7 1pa2   0 0.1 1.5   CA   308 THRA 170 ....... |         1.
 4 1  OSDOTOD   Ca    0   17    2    0    3    2   -1 ......7 1aru   0 0.2 1.6   CA   347 SER   185 ..b.... |         1.
 5 1  OTODS      Ca    0    2    2    2   -1   -1   -1 ......7 1d2v   0 0.2 1.8   CA   600 THRC 168 ....... |         1.
 6 2  OOTOO      Ca    2    0  188    5   -1   -1   -1 ......6 1jg8   0 0.5 1.8   CA   904 THRC   12 ......   |         4.
 7 3  OOOT       Ca    2    3    0   -1   -1   -1   -1 ....   5 1scj   0 0.4 2.0   CA   382 ALAA 169 ....     |    Z    3.
 8 3  ODOSOO    Ca    0    3    0    2   30   -1   -1 ......6 1e8u   0 0.2 2.0   CA 1003 ASPB 261 ......   |         3.
 9 1  OTDOND    Ca    0    2   76    3    1   -1   -1 ......7 1nbc   0 0.2 1.8   CA     1 THRB   44 ..b...   |       Z
10 1  OSEEE      Ca    0    2    4   81   -1   -1   -1 ..... 6 1iod   0 0.3 2.3   CA   501 SERA   41 ...bb    |     Z   3.
11 4  OTDOTOD   Ca    0   51    3    0    3    2   -1 ......7 1pa2   0 0.1 1.5   CA   308 THRA 170 ....... |         1.
12 4  OSDOTOD   Ca    0   17    2    0    3    2   -1 ......7 1aru   0 0.2 1.6   CA   347 SER   185 ..b.... |         1.
13 1  OS         Ca    0   -1   -1   -1   -1   -1   -1 ..     7 1edm   0 0.2 1.5   CA     3 SERB   53 ..        | ZZZZZ
```

In 5 and 6 there were originally -99 in cngp definition, i.e. donor atoms come from two different protein chains. in.tors adjusted so that required donor pair OT is picked up.

for relseq = 0

group 1 pc 1-4   (rms 3-10 deg)   3T and 1S

```
      relseq      meanfi            meanpsi        number
      0         -68(   8)         -11(   5)           4         k
```

local confs all different

group 2   pc 5-12   6T and 3S

```
            pc 5 6     (rms 12)    seems to be a continuous range of conformation
                  7                    from 5,6 through to 12.  beginning and end members
```

```
                   8 9                 up to 70 deg apart

                  10 11 12

           local confs all different


         relseq   meanfi         meanpsi     number

            0       -99( 27)     118( 30)        8      b    for whole set 5-12

            0       -63( 12)     135(  2)        2      b    for pc 5 and 6

            0     -129(  3)       78( 12)        2      d    for pc 11 and 12








      pc 13  completely different
```

```
Ca OO 2


   amino-acid sequences ****
     start of selected part of cngroup         |
     relseq                -10         -5          0          5         10         15
  1   1fzc from   312 C:   F S T W D N D N D K F E G N C A E Q D G S G
  2   3fib from   312  :   F S T W D N D N D K F E G N C A E Q D G S G
  3   1ava from   101 B:   R G I Y C I F E G G T P D A R L D W G P H M
  4   1kap from   243 P:   Q K L Y G A N L T T R T G D T V Y G F N S N
  5   1kap from   324 P:   A G V T V E N A I G G S G S D L L I G N D V
  6   1kap from   333 P:   G G S G S D L L I G N D V A N V L K G G A G
  7   1qla from   383 A:   A K L K G L F S A G E A A C W D M H G F N R
  8   1scj from   159 A:   S G S T S T V G Y P A K Y P S T I A V G A V
  9   1svy from   217  :   V A R A I D A E R K G L P K V E V _ _ _ T D
 10   1e8u from   254 A:   V T E T E E E D Y N S A V P T L M A H G R L
 11   1c8n from   163 A:   D V A P G S R V Q L S Q T Y K A I N F P P Y
 12   1dmu from    89 A:   A V G S K L T R L V S Q R E D I T V R E I G
 13   1i76 from   159 A:   I L A H A F Q P G Q G I G G D A H F D A E E
 14   1e43 from   290 A:   F N L Q A A S S Q G G G Y D M R K L L N G T
 15   1nzy from   193 A:   E V A W K V A R E L A A A P T H L Q V M A K
 16   1jg8 from     2 B:   _ _ _ _ I D L R S D T V T K P T E E M R K A
 17   1sra from   231  :   S H T E L A P L R A P L I P M E H C T T R F
 18   1ga6 from   334 A:   N G Y G G Y G Y N A G T G W D Y P T G W G S
 19   1i76 from   145 A:   R D H G D N S P F D G P N G I L A H A F Q P
 20   2por from   128  :   F S V A A S M S D G K V G E T S E D D A Q E
 21   1poc from     0  :   _ _ I Y P G T L W C G H G N K S S G P N E L
 22   1axn from    22  :   S V D A E A I Q K A I R G I G T D E K M L I
 23   1axn from   180  :   A Q I L Y K A G E N R W G T D E D K F T E I
 24   1axn from    24  :   D A E A I Q K A I R G I G T D E K M L I S I
 25   1kap from   360 P:   G G L G A D Q L W G G A G A D T F V Y G D I
 26   1kap from   351 P:   G G A G N D I L Y G G L G A D Q L W G G A G
 27   1kap from   342 P:   G N D V A N V L K G G A G N D I L Y G G L G
```

```
28   1poc from    -2  :  _ _ _ _ I Y P G T L W C G H G N K S S G P N

29   1i76 from   147 A:  H G D N S P F D G P N G I L A H A F Q P G Q

30   1cru from   259 B:  K G G N Y G W P N V A G Y K D D S G Y A Y A

31   1kap from   341 P:  I G N D V A N V L K G G A G N D I L Y G G L

32   1kap from   359 P:  Y G G L G A D Q L W G G A G A D T F V Y G D

33   1kap from   350 P:  K G G A G N D I L Y G G L G A D Q L W G G A

34   1cse from    69 E:  A G T V A A L D N T T G V L G V A P S V S L

35   1gci from    69  :  A G T I A A L N N S I G V L G V A P S A E L

36   1scj from    69 A:  A G T I A A L N N S I G V L G V S P S A S L

37   2sic from    69 E:  A G T V A A L N N S I G V L G V A P S A S L

38   1b0p from 1049 B:  F L K V L K E A E S F P G P S L V I A Y A T

39   1pa2 from    36 A:  L I R L H F H D C F V N G C D A S I L L D D

40   1jg8 from   226 B:  A R K M L G G G M R Q A G V L A A A G I I A


    aa types - acidic, etc **

     relseq                -10         -5          0          5         10         15

 1   1fzc from   312 C:  h s s h a p a p a b h a g p h s a p a g s g

 2   3fib from   312  :  h s s h a p a p a b h a g p h s a p a g s g

 3   1ava from   101 B:  b g h h h h h a g g s h a s b h a h g h p h

 4   1kap from   243 P:  p b h h g s p h s s b s g a s h h g h p s p

 5   1kap from   324 P:  s g h s h a p s h g g s g s a h h h g p a h

 6   1kap from   333 P:  g g s g s a h h h g p a h s p h h b g g s g

 7   1qla from   383 A:  s b h b g h h s s g a s s h h a h p g h p b

 8   1scj from   159 A:  s g s s s s h g h h s b h h s s h s h g s h

 9   1svy from   217  :  h s b s h a s a b b g h h b h a h _ _ _ s a

10   1e8u from   254 A:  h s a s a a a a h p s s h h s h h s p g b h

11   1c8n from   163 A:  a h s h g s b h p h s p s h b s h p h h h h

12   1dmu from    89 A:  s h g s b h s b h h s p b a a h s h b a h g

13   1i76 from   159 A:  h h s p s h p h g p g h g g a s p h a s a a

14   1e43 from   290 A:  h p h p s s s s p g g g h a h b b h h p g s

15   1nzy from   193 A:  a h s h b h s b a h s s s h s p h p h h s b

16   1jg8 from     2 B:  _ _ _ _ h a h b s a s h s b h s a a h b b s

17   1sra from   231  :  s p s a h s h h b s h h h h h a p h s s b h
```

```
18  1ga6 from  334 A:  p g h g g h g h p s g s g h a h h s g h g s
19  1i76 from  145 A:  b a p g a p s h h a g h p g h h s p s h p h
20  2por from  128  :  h s h s s s h s a g b h g a s s a a a s p a
21  1poc from    0  :  _ _ h h h g s h h h g p g p b s s g h p a h
22  1axn from   22  :  s h a s a s h p b s h b g h g s a a b h h h
23  1axn from  180  :  s p h h h b s g a p b h g s a a a b h s a h
24  1axn from   24  :  a s a s h p b s h b g h g s a a b h h h s h
25  1kap from  360 P:  g g h g s a p h h g g s g s a s h h h g a h
26  1kap from  351 P:  g g s g p a h h h g g h g s a p h h g g s g
27  1kap from  342 P:  g p a h s p h h b g g s g p a h h h g g h g
28  1poc from   -2  :  _ _ _ _ h h h g s h h h g p g p b s s g h p
29  1i76 from  147 A:  p g a p s h h a g h p g h h s p s h p h g p
30  1cru from  259 B:  b g g p h g h h p h s g h b a a s g h s h s
31  1kap from  341 P:  h g p a h s p h h b g g s g p a h h h g g h
32  1kap from  359 P:  h g g h g s a p h h g g s g s a s h h h g a
33  1kap from  350 P:  b g g s g p a h h h g g h g s a p h h g g s
34  1cse from   69 E:  s g s h s s h a p s s g h h g h s h s h s h
35  1gci from   69  :  s g s h s s h p p s h g h h g h s h s s a h
36  1scj from   69 A:  s g s h s s h p p s h g h h g h s h s s s h
37  2sic from   69 E:  s g s h s s h p p s h g h h g h s h s s s h
38  1b0p from 1049 B:  h h b h h b a s a s h h g h s h h h s h s s
39  1pa2 from   36 A:  h h b h p h p a h h h p g h a s s h h h a a
40  1jg8 from  226 B:  s b b h h g g g h b p s g h h s s s g h h s


  conformation sequences **
   relseq              -10         -5          0          5         10         15
 1  1fzc from  312 C:  b b b a k . b d b k b a . b a a a a k g b b
 2  3fib from  312  :  b b b a k b b k b k b a . b a a a a a g . b
 3  1ava from  101 B:  k g b b k b b b k g b k b k b g d b b k a k
 4  1kap from  243 P:  a a a k j b b k k b g k b b b b b . b g b k
 5  1kap from  324 P:  b g b b b a b b b b . k b b b b b b b b k b
 6  1kap from  333 P:  b . k b b b b b b b b k b b b b b b a j b j
 7  1qla from  383 A:  b a b b g b b b b j a k b d a g k a . k . b
```

```
 8   1scj from   159 A:   b . k b b k b b b k a a b a a b b b b b b b
 9   1svy from   217  :   a a a a a a a a k g g a b b b b b . . . k b
10   1e8u from   254 A:   b k b b a a a a a k b a b b b b b b b b b b
11   1c8n from   163 A:   b b b b k . a a a a a k b b g b b b . b k k
12   1dmu from    89 A:   a a a a a a a a a a a k b a k b b b b b b b
13   1i76 from   159 A:   b a . b b b b b . b . a b g b b b b b k k b
14   1e43 from   290 A:   a a a a a a k k k g g k b b k k k a k b g b
15   1nzy from   193 A:   a a a a a a a a a a a b b a a a a a a a a a
16   1jg8 from     2 B:   . . . . b b k a . k a k b b b b a a a a a a
17   1sra from   231  :   b a a k k a a k k b a k b b g k k k a a a a
18   1ga6 from   334 A:   d . b g g b j b b b . b j b b a a k j k j b
19   1i76 from   145 A:   . b . g k g d b b k j b j j b a . b b b b b
20   2por from   128  :   b b b b b b . b k g b b . d b k b b b b b b
21   1poc from     0  :   . . b b b g b k k b j b j b d b k b a k b b
22   1axn from    22  :   a a a a a a a a a a a b j b j b b a a a a a
23   1axn from   180  :   a a a a a a k . a k b b j b b a a a a a a a
24   1axn from    24  :   a a a a a a a a a b j b j b b a a a a a a a
25   1kap from   360 P:   a j b j b b b b b a j b j b b b b b b k b a
26   1kap from   351 P:   a j b j b b b b b a j b j b b b b b a j b j
27   1kap from   342 P:   b b k b b b b b b a j b j b b b b b a j b j
28   1poc from    -2  :   . . . . b b b g b k k b j b j b d b k b a k
29   1i76 from   147 A:   . g k g d b b k j b j j b a . b b b b b . b
30   1cru from   259 B:   b g b b k g b k a b b j b b k g b g b b b b
31   1kap from   341 P:   b b b k b b b b b b a j b j b b b b b a j b
32   1kap from   359 P:   b a j b j b b b b b a j b j b b b b b b k b
33   1kap from   350 P:   b a j b j b b b b b a j b j b b b b b a j b
34   1cse from    69 E:   a a a a k b b b . k b j b b b a b k k b b b
35   1gci from    69  :   a a a a . b b b . k b j b b b a d k k b b b
36   1scj from    69 A:   a a a a d b b b . k b . b b b a b k k b b b
37   2sic from    69 E:   a a a a d b b b . d b j b b b a d a k b b b
38   1b0p from  1049 B:   a a a a a a a a a k b d . b b b b b b b b b
39   1pa2 from    36 A:   a a a a a a a a a k b g j d k g a a k b b b
40   1jg8 from   226 B:   a a a a a g k b b a d k a a a a a a a a a a
```

```
 **********  protein names (pdb header) *******

 1  1fzc at  322 C   2.3      BLOOD COAGULATION   Fibrinogen C-terminal domain-like unusual fold

 2  3fib at  322     2.1      BLOOD COAGULATION    Fibrinogen C-terminal domain-like unusual fold

 3  1ava at  111 B   1.9      HYDROLASE INHIBITION   beta-Trefoil   Amylase/subtilisin inhibitor

 4  1kap at  253 P   1.64     ZINC METALLOPROTEASE beta-Roll  Metalloprotease, C-terminal d

 5  1kap at  334 P   1.64     ZINC METALLOPROTEASE

 6  1kap at  343 P   1.64     ZINC METALLOPROTEASE

 7  1qla at  393 A   2.2      OXIDOREDUCTASE       FAD/NAD(P)-binding domain Succinate dehydrogenase/fumara

 8  1scj at  169 A   2.0      HYDROLASE          Subtilisin-like

 9  1svy at  227      1.75     ACTIN-BINDING PROTEIN

10  1e8u at  264 A   2.0      SIALIDASE

11  1c8n at  173 A   2.25     VIRUS/VIRAL PROTEIN

12  1dmu at   99 A   2.2      HYDROLASE/DNA

13  1i76 at  169 A   1.20     HYDROLASE

14  1e43 at  300 A   1.7      HYDROLASE

15  1nzy at  203 A   1.8      LYASE

16  1jg8 at   12 B   1.80     LYASE

17  1sra at  241     2.0      CALCIUM-BINDING PROTEIN


18  1ga6 at  344 A   1.00     HYDROLASE    : Subtilisin-like  Serine-carboxyl proteinase

19  1i76 at  155 A   1.20     HYDROLASE     Zincin-like    Matrix metalloproteases, catalyti

20  2por at  138     1.8      INTEGRAL MEMBRANE PROTEIN PORIN  Membrane and cell surface proteins and peptides

21  1poc at   10     2.0      HYDROLASE   Phospholipase A2, PLA2

22  1axn at   32      1.78     CALCIUM/PHOSPHOLIPID-BINDING PROTEIN   Annexin

23  1axn at  190     1.78     CALCIUM/PHOSPHOLIPID-BINDING PROTEIN

24  1axn at   34      1.78     CALCIUM/PHOSPHOLIPID-BINDING PROTEIN

25  1kap at  370 P   1.64     ZINC METALLOPROTEASE   beta-Roll

26  1kap at  361 P   1.64     ZINC METALLOPROTEASE

27  1kap at  352 P   1.64     ZINC METALLOPROTEASE

28  1poc at    8      2.0      HYDROLASE


29  1i76 at  157 A   1.20     HYDROLASE              Zincin-like
```

```
30  1cru at  269 B   1.50     OXIDOREDUCTASE         6-bladed beta-propeller
31  1kap at  351 P   1.64     ZINC METALLOPROTEASE       beta-Roll
32  1kap at  369 P   1.64     ZINC METALLOPROTEASE
33  1kap at  360 P   1.64     ZINC METALLOPROTEASE
34  1cse at   79 E   1.2       COMPLEX(SERINE PROTEINASE-INHIBITOR)Subtilisin-like
35  1gci at   79      0.78     SUBTILISIN FROM BACILLUS LENTUS       Subtilisin-like
36  1scj at   79 A   2.0       HYDROLASE                                 Subtilisin-like
37  2sic at   79 E   1.8       COMPLEX (PROTEINASE/INHIBITOR)        Subtilisin-like
38  1b0p at 1059 B   2.31     OXIDOREDUCTASE                Thiamin diphosphate-binding fold (THDP-binding)
39  1pa2 at   46 A   1.45     OXIDOREDUCTASE                  Heme-dependent peroxidases
40  1jg8 at  236 B   1.80     LYASE                                PLP-dependent transferases


 **********  cngps  ************
 1 3  DDOO      Ca   2   2   2  -1  -1  -1  -1 ....  4 1fzc  0 0.2 2.3  CA    1 ASPC 318 b...     |
 2 3  DDOO      Ca   2   2   2  -1  -1  -1  -1 ....  6 3fib  0 0.2 2.1  CA  400 ASP  318 b...     |   ZZ
 3 2  EOOD      Ca   3   2   4  -1  -1  -1  -1 ....  5 1ava  0 0.4 1.9  CA  501 GLUB 108 b..b     |   Z   3.
 4 1  OOTDOD   Ca   2   2  28   2   3  -1  -1 ......6 1kap  1 0.1 1.6  CA  614 ARGP 253 ...b..  |        3.
 5 1  OODOOD   Ca   2   2  13   2   3  -1  -1 ......6 1kap  0 0.3 1.6  CA  616 GLYP 334 .....b  |        3.
 6 1  OONOOD   Ca   2   2  13   2   3  -1  -1 ......6 1kap  0 0.3 1.6  CA  618 ASNP 343 .....b  |        3.
 7 4  OOOOO     Ca   1   1  20   2  -1  -1  -1 ..... 6 1qla  0 0.2 2.2  CA    9 SERA 371 .....    |    Z  1.
 8 1  OOOT      Ca   2   3   0  -1  -1  -1  -1 ....  5 1scj  0 0.4 2.0  CA  382 ALAA 169 ....     |   Z   3.
 9 2  DOO        Ca   5   2  -1  -1  -1  -1  -1 ...    6 1svy  0 0.3 1.8  CA    1 ASP  222 ...      |  ZZZ
10 3  ODOOO     Ca   0   3   2  30  -1  -1  -1 ..... 5 1e8u  0 0.1 2.0  CA 1002 ASPA 261 .....    |        3.
11 1  OO         Ca   2  -1  -1  -1  -1  -1  -1 ..     2 1c8n  0 0.5 2.3  CA  305 SERA 173 ..        |
12 1  OOO        Ca   2   3  -1  -1  -1  -1  -1 ...    6 1dmu  0 0.2 2.2  CA  125 SERA  99 ...      |  ZZZ  3.
13 2  OOOD      Ca  32   2   2  -1  -1  -1  -1 ....  6 1i76  0 0.1 1.2  CA  996 ASPA 137 ....     |   ZZ  3.
14 1  OOODD     Ca   2 104   1  23  -1  -1  -1 .. .. 6 1e43  0 0.1 1.7  CA  504 GLYA 300 ....b    |    Z  3.
15 3  OOOOTQ   Ca 153   1   2   2   3  -1  -1 ......6 1nzy  0 0.4 1.8  CA  271 GLYA  49 ......  |        3.
16 1  OOTOOQ   Ca   2   0 188   5 -99  -1  -1 ......6 1jg8  0 0.5 1.8  CA  901 THRB  12 ......  |        4.
17 1  OOE        Ca   2   3  -1  -1  -1  -1  -1 ...    5 1sra  1 0.4 2.0  CA  303 PRO  241 ...      |  ZZ


18 3  DOOOD     Ca   1  15   2   2  -1  -1  -1 ..... 6 1ga6  0 0.1 1.0  CA  374 ASPA 328 .....    |    Z  3.
19 2  DOOODE   Ca   1   2   2  18   3  -1  -1 ......6 1i76  0 0.1 1.2  CA  997 ASPA 154 ......  |        3.
```

```
20 3  NDOO      Ca  20   2   2  -1  -1  -1  -1 ....  5 2por  0 0.2 1.8  CA  304 ASN  116 .b..    |   Z

21 2  OOOD      Ca   2   2  23  -1  -1  -1  -1 ....  6 1poc  0 0.1 2.0  CA  501 TRP    8 ...b    |   UU  3.

22 1  OOOD      Ca   2   2  40  -1  -1  -1  -1 ....  4 1axn  0 0.4 1.8  CA  351 ILE   32 ...b    |

23 2  OOOE      Ca   3   2  40  -1  -1  -1  -1 ....  6 1axn  0 0.2 1.8  CA  353 GLY  187 ...b    |   ZZ

24 2  OOOD      Ca   2   2  40  -1  -1  -1  -1 ....  4 1axn  0 0.4 1.8  CA  351 ILE   32 ...b    |

25 1  OODD      Ca   2   2  26  -1  -1  -1  -1 ....  6 1kap  1 0.2 1.6  CA  617 GLYP 370 ....    |   ZZ  3.

26 1  OODOD     Ca   2   2  18   7  -1  -1  -1 ..... 6 1kap  0 0.1 1.6  CA  619 GLYP 361 .....   |    Z  3.

27 1  OODOOD   Ca   2   2  13   2   3  -1  -1 ......6 1kap  0 0.3 1.6  CA  620 GLYP 352 .....b  |        3.

28 1  OOOD      Ca   2   2  23  -1  -1  -1  -1 ....  6 1poc  0 0.1 2.0  CA  501 TRP    8 ...b    |   UU  3.


29 3  DOOODE   Ca   1   2   2  18   3  -1  -1 ......6 1i76  0 0.1 1.2  CA  997 ASPA 154 ......  |        3.

30 1  OODE      Ca   2   2  36  -1  -1  -1  -1 ....  6 1cru  0 0.1 1.5  CA  902 ALAB 269 ...b    |   ZZ  1.

31 4  OODOOD   Ca   2   2  13   2   3  -1  -1 ......6 1kap  0 0.3 1.6  CA  616 GLYP 334 .....b  |        3.

32 4  OODOOD   Ca   2   2  13   2   3  -1  -1 ......6 1kap  0 0.3 1.6  CA  620 GLYP 352 .....b  |        3.

33 4  OONOOD   Ca   2   2  13   2   3  -1  -1 ......6 1kap  0 0.3 1.6  CA  618 ASNP 343 .....b  |        3.

34 5  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1cse  0 0.1 1.2  CA  430 GLNE   2 .b....  |        3.

35 5  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1gci  0 0.1 0.8  CA  277 GLN    2 .b....  |        3.

36 5  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1scj  0 0.2 2.0  CA  381 GLNA   2 .b....  |        3.

37 5  QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 2sic  0 0.2 1.8  CA  501 GLNE   2 .b....  |        3.

38 4  DNOOO     Ca   2  71   3   2  -1  -1  -1 ..... 5 1b0p  1 0.4 2.3  CA 1238 ASPB 983 .....   |        1.

39 3  ODOODS   Ca   0   3   2   2   2  -1  -1 ......7 1pa2  0 0.1 1.5  CA  307 ASPA  43 ......  |      Z 1.

40 2  QOOTOO   Ca -99   2   0 188   5  -1  -1 ......6 1jg8  0 0.5 1.8  CA  904 GLNB 236 ......  |        4.


NB  pc 40 incorrectly interpreted because of -99   - two protein chains

     contribute to cngp


group 1:  1-17

==============

 relseq      meanfi            meanpsi       number

      1         -70(  9)         -18( 14)          17         k

      2         -77( 15)       -198( 24)          17         b


   sub group   a)  1-7       model: 1kap at  253 P   1.64  (4)
```

```
      relseq      meanfi           meanpsi       number

          1        -69(   6)        -18(  12)          7         k

          2        -91(   7)      -173(  13)          7         b

    sub group    b)   8-10

    relseq      meanfi           meanpsi       number

          1        -58(   3)        -37(   3)          3         a

          2        -66(   5)        146(   9)          3         b

    sub group     c)   11-16     model: 1i76 at  169 A    1.20    (13)

    relseq      meanfi           meanpsi       number

          1        -74(   7)        -14(   9)          6         k

          2        -64(  10)       145(  10)          6         b


        no similarity in local conformations (-10 to 11) except that

          1 and 2 are very close    (rms 9)



  group 2: 18-28

  ==============

   relseq      meanfi           meanpsi       number

          1        -63(   8)        142(  16)         10         b

          2          76(  14)      -166(  22)         10         j


    sub group a) 18-23        model: 1i76 at  155 A    1.20     (19)

    relseq      meanfi           meanpsi       number

          1        -59(   6)        132(   9)          6         b

          2          67(  10)      -151(  11)          6         j

    sub group b) 25-27        model: 1kap at  370 P    1.64     (25)

    relseq      meanfi           meanpsi       number

          1        -65(   1)        164(   3)          3         b

          2          92(   3)        166(   5)          3         j


        no similarity in local conformations (-10 to 11) except that

     chains 25-27 are v similar over relseq 4-17
```

```
group 3: 29-37
===============
                     model 1gci at   79      0.78      (35)
  relseq      meanfi           meanpsi       number
      1          67(  7)       -141( 15)          9         j
      2        -90( 27)         173( 22)          9         b

  sub group a) 29-33
  relseq      meanfi           meanpsi       number
      1          71(  4)       -149( 15)          5         j
      2        -67(  3)         157( 14)          5         b
  sub group b) 34-37
 relseq      meanfi           meanpsi       number
      1          62(  5)       -131(  3)          4         j
      2       -118(  2)       -165(  4)          4         b

      subgroup a) local conformations different, all proteases
      subgroup b)  local conformations all same,  all proteases



+ 4 outliers   (reduced to 3 when pc 40 removed, althoug
          it may belong toone of the main groups.)
```

```
Ca OO 3


   amino-acid sequences ****

     start of selected part of cngroup         |

     relseq                -10         -5          0          5         10         15

  1   1acc from   212  :   E K K G L T K Y K S S P E K W S T A S D P Y S

  2   1sbw from    62 A:   G I Q V R _ _ _ E D N I N V V E G N E Q F I S

  3   2btc from    62 E:   G I Q V R _ _ _ E D N I N V V E G N E Q F I S

  4   1avw from    62 A:   R I Q V R _ _ _ E H N I D V L E G N E Q F I N

  5   1pyt from   462 C:   R T Y Q V V L G E Y D R S V L Q G S E Q V I P

  6   1tn3 from   137  :   E T E I T A Q P D G G K T E N C A V L S G A A

  7   1lpb from   177 B:   P A E P C F Q G T P E L V R L D P S D A K F V

  8   1ava from   133 A:   G T G N P D T G A D F G A A P D I D H L N L R

  9   1axn from   177  :   K Q D A Q I L Y K A G E N R W G T D E D K F T

 10   1slu from    62 B:   R I Q V _ _ _ G E H N I N V L E G N E Q F V N

 11   1dmu from    91 A:   G S K L T R L V S Q R E D I T V R E I G L P T

 12   1b0p from 1046 B:  K Q Q F L K V L K E A E S F P G P S L V I A Y

 13   1g5c from   108 F:   G V E E E V I E N F S I D V L N P V G D E E E

 14   1scj from   161 A:   S T S T V G Y P A K Y P S T I A V G A V N S S

 15   1a0t from   444 P:   T S W M D W S K K L N N Y A S D D A L G S D G

 16   1kit from   243  :   E K R V G G G D P G A L S N T N D I I T R T S

 17   1dx5 from   430 J:   G G F C S G V C H N L P G T F E C I C G P D S

 18   1fjs from    62 A:   K R F K V R V G D R N T E Q E E G G E A V H E

 19   1dyo from    29 A:   E V L L S G R T A Y K G S E S L L V R N R T A

 20   1fo4 from   857 A:   V A L E V D H Y S N A G N S R D L S H S I M E

 21   1brw from   233 A:   M A V I S D M S Q P L G Y A V G N A L E V K E


   aa types - acidic, etc **

     relseq                -10         -5          0          5         10         15

  1   1acc from   212  :   a b b g h s b h b s s h a b h s s s s a h h s

  2   1sbw from    62 A:   g h p h b _ _ _ a a p h p h h a g p a p h h s

  3   2btc from    62 E:   g h p h b _ _ _ a a p h p h h a g p a p h h s
```

```
 4  1avw from   62 A:  b h p h b _ _ _ a p p h a h h a g p a p h h p

 5  1pyt from  462 C:  b s h p h h h g a h a b s h h p g s a p h h h

 6  1tn3 from  137  :  a s a h s s p h a g g b s a p h s h h s g s s

 7  1lpb from  177 B:  h s a h h h p g s h a h h b h a h s a s b h h

 8  1ava from  133 A:  g s g p h a s g s a h g s s h a h a p h p h b

 9  1axn from  177  :  b p a s p h h h b s g a p b h g s a a a b h s

10  1slu from   62 B:  b h p h _ _ _ g a p p h p h h a g p a p h h p

11  1dmu from   91 A:  g s b h s b h h s p b a a h s h b a h g h h s

12  1b0p from 1046 B:  b p p h h b h h b a s a s h h g h s h h h s h

13  1g5c from  108 F:  g h a a a h h a p h s h a h h p h h g a a a a

14  1scj from  161 A:  s s s s h g h h s b h h s s h s h g s h p s s

15  1a0t from  444 P:  s s h h a h s b b h p p h s s a a s h g s a g

16  1kit from  243  :  a b b h g g g a h g s h s p s p a h h s b s s

17  1dx5 from  430 J:  g g h h s g h h p p h h g s h a h h h g h a s

18  1fjs from   62 A:  b b h b h b h g a b p s a p a a g g a s h p a

19  1dyo from   29 A:  a h h h s g b s s h b g s a s h h h b p b s s

20  1fo4 from  857 A:  h s h a h a p h s p s g p s b a h s p s h h a

21  1brw from  233 A:  h s h h s a h s p h h g h s h g p s h a h b a


  conformation sequences **

   relseq             -10        -5         0         5        10        15

 1  1acc from  212  :  a k k g b b b b b b b k k b a b a k g b k b b

 2  1sbw from   62 A:  g b b b b . . . b . b k k b b k . g b b b b b

 3  2btc from   62 E:  g b b b b . . . b a b k k b b k g k b b b b b

 4  1avw from   62 A:  b b b b b . . . b a b k k b b k g k b b b b b

 5  1pyt from  462 C:  . b b b b b b g b . b k k b b k g a b b b b b

 6  1tn3 from  137  :  b a d . b b b b k k . k k b k b b b b b a a k

 7  1lpb from  177 B:  b b a b g . b g b b a a k b b . a k k b a b b

 8  1ava from  133 A:  g k b b b b k . b b b a k b b b b b k k b a a

 9  1axn from  177  :  a a a a a a a a a k . a k b b j b b a a a a a

10  1slu from   62 B:  b b b b . . . g b . b a k b b k . b . g b b b

11  1dmu from   91 A:  a a a a a a a a a k b a k b b b b b b b b k b

12  1b0p from 1046 B:  a a a a a a a a a a a a k b d . b b b b b b b
```

```
13  1g5c from  108 F:  g b b a a a a a a a a a a b b b d a b b a a a
14  1scj from  161 A:  k b b k b b b k a a b a a b b b b b b b b k k
15  1a0t from  444 P:  b b b b b . b a a k a k a b a k b a b g b b g
16  1kit from  243   :  b b b b . j j b b g b k d . b b b b b b b b b
17  1dx5 from  430 J:  j g a b a . b b b b b b g . b b b b b . a . .
18  1fjs from   62 A:  a b b b b b b g b k b a d . b a g k b b b b b
19  1dyo from   29 A:  b b . b b k a b . a b j k b b b b b b g b k b
20  1fo4 from  857 A:  a b b b b b b b b b a j k . a . k a a a a a a
21  1brw from  233 A:  b b b b b b k a b b b j a b b . . a a a a a a


 **********  protein names (pdb header) *******
 1  1acc at  222      2.1       TOXIN   Anthrax protective antigen
 2  1sbw at   72 A    1.80      HYDROLASE/HYDROLASE INHIBITOR   Trypsin-like serine proteases
 3  2btc at   72 E    1.50      HYDROLASE/HYDROLASE INHIBITOR   Trypsin-like serine proteases
 4  1avw at   72 A    1.75      COMPLEX (PROTEINASE/INHIBITOR)  beta-Trefoil     (is trypsin inhibitor)
 5  1pyt at  472 C    2.35      TERNARY COMPLEX (ZYMOGEN)
 6  1tn3 at  147      2.0       LECTIN                                 C-type lectin-like
 7  1lpb at  187 B    2.46      HYDROLASE(CARBOXYLIC ESTERASE)     Trypsin-like serine proteases
 8  1ava at  143 A    1.9       HYDROLASE INHIBITION                TIM beta/alpha-barrel
 9  1axn at  187      1.78      CALCIUM/PHOSPHOLIPID-BINDING PROTEIN  Annexin
10  1slu at   72 B    1.8       COMPLEX (SERINE PROTEASE/INHIBITOR)  Trypsin-like serine proteases
11  1dmu at  101 A    2.2       HYDROLASE/DNA      Restriction endonuclease-like
12  1b0p at 1056 B    2.31      OXIDOREDUCTASE
13  1g5c at  118 F    2.10      LYASE
14  1scj at  171 A    2.0       HYDROLASE
15  1a0t at  454 P    2.4       OUTER MEMBRANE PROTEIN
16  1kit at  253      2.3       HYDROLASE
17  1dx5 at  440 J    2.3       SERINE PROTEINASE
18  1fjs at   72 A    1.92      BLOOD CLOTTING
19  1dyo at   39 A    2.1       CARBOHYDRATE-BINDING MODULE
20  1fo4 at  867 A    2.10      OXIDOREDUCTASE
21  1brw at  243 A    2.1       NUCLEOSIDE PHOSPHORYLASE
```

```
 **********  cngps  ************

 1 4  DDEOOD   Ca   2   7  34   3  10  -1  -1 ......6 1acc  0 0.3 2.1  CA  801 ASP  179 ..b...  |

 2 2  EOOE      Ca   2   3   5  -1  -1  -1  -1 ....  6 1sbw  0 0.2 1.8  CA  801 GLUA  70 ....    |   ZZ  3.

 3 2  EOOE      Ca   2   3   5  -1  -1  -1  -1 ....  6 2btc  0 0.2 1.5  CA  700 GLUE  70 ....    |   ZZ  3.

 4 2  EOOEE    Ca   2   3   2   3  -1  -1  -1 ..... 6 1avw  0 0.1 1.8  CA  700 GLUA  70 .....   |    Z  3.

 5 2  EOOQE    Ca   2   3   2   3  -1  -1  -1 ..... 5 1pyt  1 0.3 2.3  CA  650 GLUC 470 .....   |        3.

 6 3  DEOON    Ca   4  27   3   1  -1  -1  -1 ..... 6 1tn3  0 0.2 2.0  CA  182 ASP  116 bb...   |    Z

 7 1  OODD      Ca   3   2   3  -1  -1  -1  -1 ....  6 1lpb  0 0.2 2.5  CA  450 GLUB 187 ...b    |   ZZ  3.

 8 3  DDOOD    Ca  15   1   3   2  -1  -1  -1 ..... 6 1ava  1 0.2 1.9  CA  502 ASPA 127 .....   |    Z  3.

 9 1  OOOE      Ca   3   2  40  -1  -1  -1  -1 ....  6 1axn  0 0.2 1.8  CA  353 GLY  187 ...b    |   ZZ

10 2  EOOEO    Ca   2   3   2 -99  -1  -1  -1 ..... 7 1slu  0 0.5 1.8  CA  101 GLUB  70 .....   |    UZ 3.

11 2  OOO       Ca   2   3  -1  -1  -1  -1  -1 ...   6 1dmu  0 0.2 2.2  CA  125 SERA  99 ...     |  ZZZ  3.

12 3  DNOOO    Ca   2  71   3   2  -1  -1  -1 ..... 5 1b0p  1 0.4 2.3  CA 1238 ASPB 983 .....   |        1.

13 2  EOO       Ca  21   3  -1  -1  -1  -1  -1 ...   6 1g5c  0 0.3 2.1  CA 1009 GLUF  97 b..     |  ZZZ  4.

14 2  OOOT      Ca   2   3   0  -1  -1  -1  -1 ....  5 1scj  0 0.4 2.0  CA  382 ALAA 169 ....    |   Z   3.

15 1  OOO       Ca   3   5  -1  -1  -1  -1  -1 ...   3 1a0t  0 0.1 2.4  CA   10 ASNP 454 ...     |

16 1  OONDOT   Ca   3   0  33  24   0  -1  -1 ......6 1kit  0 0.2 2.3  CA  802 ALA  253 ...b..  |        3.

17 5  DOENOO   Ca   1   2  13   1   3  -1  -1 ......7 1dx5  0 0.2 2.3  CA 1001 ASPJ 423 b.....  |     Z 3.

18 2  DOOEE    Ca   2   3   2   3  -1  -1  -1 ..... 6 1fjs  0 0.3 1.9  CA  507 ASPA  70 .....   |    Z  3.

19 3  OEOOD    Ca   2  23   3 107  -1  -1  -1 ..... 5 1dyo  0 0.4 2.1  CA  300 THRA  14 ....b   |

20 1  OOOSSO   Ca   3   1   3  33   1  -1  -1 ......6 1fo4  0 0.4 2.1  CA 4009 ALAA 867 ......  |        1.

21 3  OTOOE    Ca   2 153   3   9  -1  -1  -1 ..... 5 1brw  0 0.3 2.1  CA 3001 GLYA  88 .....   |        2.



for chelate loop use fi and psi 1-3



group 1: pc 1-6       model: 2btc at   72 E   1.50     (3)

  relseq     meanfi          meanpsi      number

      1        -72(  4)         -4(  5)          6        k

      2       -101(  6)         -8(  7)          6        k

      3       -129( 12)        127( 12)          6        b

  local conformations 3 4 5 similar
```

```
group 2:  pc 7-10     model: 1slu at   72 B   1.8   (10)

  relseq     meanfi            meanpsi       number

     1         -58(  5)         -40(  4)           4         a

     2         -79(  7)           -5( 11)           4         k

     3       -144(  9)        163( 11)           4         b

  local conformations all different



group 3:  pc 11-14     model: 1g5c at  118 F   2.10  (13)

   relseq     meanfi            meanpsi       number

     1         -54(  9)         -45( 10)           4         a

     2         -72( 10)         -19( 18)           4         a

     3         -83( 20)         131( 26)           4         b

 local conf: 11,14 quite b at relseq 0, 12 13 are a, and other difs


groups 1-3, and pc 15 16 are all within ca 45 deg r.m.s

   relseq     meanfi            meanpsi       number

     1         -64( 11)         -24( 20)          16         a

     2         -85( 17)           -9( 21)          16         k

     3       -125( 30)        142( 23)          16         b


  pc 17-21 are quite different from them, and mostly from each other,

           except that 18 19 21 are within about 30 deg of each other

   for 19 21  relseq     meanfi            meanpsi       number

                 1          67(  8)       -149(  1)           2         j

                 2       -116(  1)         -17( 27)           2         k

                 3       -103( 30)         154( 21)           2         b
```

```
Zn H H 4
  amino-acid sequences ****
   start of selected part of cngroup         |
   relseq                 -10         -5           0           5          10          15
 1  1a1i from  143 A:  N F S R S D H L T T H I R T H T G E K P F A C D
 2  1a1i from  171 A:  K F A R S D E R K R H T K I H L R _ _ _ _ _ _ _
 3  1rmd from   98  :   E E V S L E K Y N H H V S S H K E S K _ _ _ _ _
 4  1a1i from  115 A:  R F S R S A D L T R H I R I H T G Q K P F Q C R
 5  1bs4 from 1122 C:  A D G L L A I C I Q H E M D H L V G K L F M D Y
 6  1j98 from   44 A:   N K Q A M K P D T I H T L E H L L A F T I R S H
 7  1c7k from   73 A:   Q Q Y D S T R V T A H E T G H V L G L P D H Y Q
 8  1g12 from  107 A:  G T D S Q A G T L V H E S S H F T R N G G T K D
 9  1hs6 from  285 A:  G D K S L S N V I A H E I S H S W T G N L V T N
10  1ezm from  130  :  Y P L V S L D V A A H E V S H G F T E Q N S G L
11  1i1i from  464 P:  L R H D E V R T Y F H E F G H V M H Q I C A Q T
12  1dmt from  573 A:  N Y G G I G M V I G H E I T H G F D D N G R N F
13  1hr6 from   60 B:  A E N V K N N G T A H F L E H L A F K G T Q N R
14  1i76 from  187 A:  A N Y N L F L V A A H E F G H S L G L A H S S D
15  1lml from  254  :  Y D Q L V T R V V T H E M A H A L G F S G P F F
16  1bkc from  395 E:  L T K E A D L V T T H E L G H N F G A E H D P D
17  1kap from  166 P:  N G N Y G R Q T L T H E I G H T L G L S H P G D
18  1atl from  132 A:  I N L L M G V T M A H E L G H N L G M E H D G K

  aa types - acidic, etc **
   relseq                 -10         -5           0           5          10          15
 1  1a1i from  143 A:  p h s b s a p h s s p h b s p s g a b h h s h a
 2  1a1i from  171 A:  b h s b s a a b b b p s b h p h b _ _ _ _ _ _ _
 3  1rmd from   98  :   a a h s h a b h p p p h s s p b a s b _ _ _ _ _
 4  1a1i from  115 A:  b h s b s s a h s b p h b h p s g p b h h p h b
 5  1bs4 from 1122 C:  s a g h h s h h h p p a h a p h h g b h h h a h
 6  1j98 from   44 A:   p b p s h b h a s h p s h a p h h s h s h b s p
 7  1c7k from   73 A:   p p h a s s b h s s p a s g p h h g h h a p h p
```

```
 8   1g12  from   107  A:   g  s  a  s  p  s  g  s  h  h  p  a  s  s  p  h  s  b  p  g  g  s  b  a
 9   1hs6  from   285  A:   g  a  b  s  h  s  p  h  h  s  p  a  h  s  p  s  h  s  g  p  h  h  s  p
10   1ezm  from   130   :   h  h  h  h  s  h  a  h  s  s  p  a  h  s  p  g  h  s  a  p  p  s  g  h
11   1i1i  from   464  P:   h  b  p  a  a  h  b  s  h  h  p  a  h  g  p  h  h  p  p  h  h  s  p  s
12   1dmt  from   573  A:   p  h  g  g  h  g  h  h  h  g  p  a  h  s  p  g  h  a  a  p  g  b  p  h
13   1hr6  from    60  B:   s  a  p  h  b  p  p  g  s  s  p  h  h  a  p  h  s  h  b  g  s  p  p  b
14   1i76  from   187  A:   s  p  h  p  h  h  h  h  s  s  p  a  h  g  p  s  h  g  h  s  p  s  s  a
15   1lml  from   254   :   h  a  p  h  h  s  b  h  h  s  p  a  h  s  p  s  h  g  h  s  g  h  h  h
16   1bkc  from   395  E:   h  s  b  a  s  a  h  h  s  s  p  a  h  g  p  p  h  g  s  a  p  a  h  a
17   1kap  from   166  P:   p  g  p  h  g  b  p  s  h  s  p  a  h  g  p  s  h  g  h  s  p  h  g  a
18   1atl  from   132  A:   h  p  h  h  h  g  h  s  h  s  p  a  h  g  p  p  h  g  h  a  p  a  g  b


   conformation  sequences  **
     relseq                 -10           -5            0             5           10           15
 1   1a1i  from   143  A:   b  b  a  b  a  a  a  a  a  a  a  a  k  a  a  a  g  b  b  k  b  b  b  a
 2   1a1i  from   171  A:   b  b  a  b  a  a  a  a  a  a  a  a  k  k  k  b  .  .  .  .  .  .  .  .
 3   1rmd  from    98   :    b  b  b  b  a  a  k  a  a  a  a  a  a  k  b  b  b  a  .  .  .  .  .  .
 4   1a1i  from   115  A:   b  b  a  b  a  a  a  a  a  a  a  a  a  a  a  k  g  b  b  k  b  b  b  k
 5   1bs4  from  1122  C:   b  b  j  a  a  a  a  a  a  a  a  a  a  a  a  k  k  g  b  b  a  k  k  k
 6   1j98  from    44  A:   g  a  b  b  b  b  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a
 7   1c7k  from    73  A:   a  a  b  b  a  a  a  a  a  a  a  a  a  a  a  a  k  g  b  b  b  d  a  k
 8   1g12  from   107  A:   j  b  g  b  a  a  a  a  a  a  a  a  a  a  a  b  a  a  k  .  g  b  b  b
 9   1hs6  from   285  A:   g  k  g  a  k  k  a  a  a  a  a  a  a  a  a  k  a  b  .  a  a  b  b  b
10   1ezm  from   130   :   b  b  k  b  b  a  a  a  a  a  a  a  a  k  a  a  a  a  a  k  .  a  g  b
11   1i1i  from   464  P:   b  b  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  k  b  b  b
12   1dmt  from   573  A:   a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  k  k  b  k  a  a  a  k  b
13   1hr6  from    60  B:   k  b  b  a  a  a  b  g  a  a  a  a  a  a  a  k  k  k  k  .  b  k  a  b
14   1i76  from   187  A:   b  b  b  b  a  a  a  a  a  a  a  a  a  a  a  a  k  g  b  b  b  b  k  b
15   1lml  from   254   :   d  b  a  a  a  a  a  a  a  a  a  a  a  a  a  a  k  .  b  b  a  a  a  a
16   1bkc  from   395  E:   b  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  k  .  b  b  b  b  b  b
17   1kap  from   166  P:   b  g  b  a  a  a  a  a  a  a  a  a  a  a  a  a  k  .  b  b  b  k  b  b
18   1atl  from   132  A:   b  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  k  g  b  b  b  b  b  k
```

```
 **********  protein names (pdb header) *******

  1  1a1i at  153 A   1.6       COMPLEX (ZINC FINGER/DNA)
  2  1a1i at  181 A   1.6       COMPLEX (ZINC FINGER/DNA)
  3  1rmd at  108     2.1       DNA-BINDING PROTEIN
  4  1a1i at  125 A   1.6       COMPLEX (ZINC FINGER/DNA)
  5  1bs4 at 1132 C   1.90      HYDROLASE
  6  1j98 at   54 A   1.20      SIGNALING PROTEIN
  7  1c7k at   83 A   1.00      HYDROLASE
  8  1g12 at  117 A   1.6       HYDROLASE
  9  1hs6 at  295 A   1.95      HYDROLASE
 10  1ezm at  140     1.5       HYDROLASE
 11  1i1i at  474 P   2.30      HYDROLASE
 12  1dmt at  583 A   2.10      HYDROLASE
 13  1hr6 at   70 B   2.50      HYDROLASE
 14  1i76 at  197 A   1.20      HYDROLASE
 15  1lml at  264     1.86      LEISHMANOLYSIN
 16  1bkc at  405 E   2.0       ZN-ENDOPEPTIDASE
 17  1kap at  176 P   1.64      ZINC METALLOPROTEASE
 18  1atl at  142 A   1.8       METALLOENDOPEPTIDASE


 **********  cngps  ************

  1 3  CCHH     Zn   3  13   4  -1  -1  -1  -1 ..ee  4 1a1i  0 0.1 1.6  ZN  202 CYSA 137 ....    |
  2 3  CCHH     Zn   3  13   4  -1  -1  -1  -1 ..ee  4 1a1i  0 0.1 1.6  ZN  203 CYSA 165 ....    |
  3 3  CCHH     Zn   5  12   4  -1  -1  -1  -1 ..ee  4 1rmd  0 0.1 2.1  ZN  120 CYS   91 ....    |
  4 3  CCHH     Zn   5  13   4  -1  -1  -1  -1 ..ee  4 1a1i  0 0.1 1.6  ZN  201 CYSA 107 ....    |
  5 2  CHH      Zn  42   4  -1  -1  -1  -1  -1 .ee   4 1bs4  0 0.1 1.9  ZN 2003 CYSC1090 ...     |  U     3.
  6 1  HHC      Zn   4  68  -1  -1  -1  -1  -1 ee.   4 1j98  0 0.1 1.2  ZN  300 HISA  54 ...     |  Z
  7 1  HHD      Zn   4   6  -1  -1  -1  -1  -1 ee.   4 1c7k  0 0.1 1.0  ZN  133 HISA  83 ...     |  Z     3.
  8 1  HHD      Zn   4   9  -1  -1  -1  -1  -1 ee.   5 1g12  0 0.2 1.6  ZN  200 HISA 117 ..b     |  ZZ    3.
  9 1  HHE      Zn   4  19  -1  -1  -1  -1  -1 ee.   5 1hs6  1 0.2 2.0  ZN  701 HISA 295 ...     |  UU    3.
 10 1  HHE      Zn   4  20  -1  -1  -1  -1  -1 ee.   4 1ezm  1 0.2 1.5  ZN  300 HIS  140 ...     |  Z     3.
 11 1  HHE      Zn   4  25  -1  -1  -1  -1  -1 ee.   4 1i1i  0 0.4 2.3  ZN  701 HISP 474 ..b     |  Z     3.
 12 1  HHE      Zn   4  59  -1  -1  -1  -1  -1 ee.   4 1dmt  0 0.4 2.1  ZN  755 HISA 583 ..b     |  U     3.
```

```
13 1   HHE        Zn    4   76   -1   -1   -1   -1   -1 ee.    4 1hr6   0 0.1 2.5   ZN   501 HISB   70 ..b       |   Z      3.
14 1   HHH        Zn    4    6   -1   -1   -1   -1   -1 eee    5 1i76   0 0.3 1.2   ZN   999 HISA 197 ...       |   UU    3.
15 1   HHH        Zn    4   66   -1   -1   -1   -1   -1 eee    4 1lml 20 0.4 1.9   ZN   100 HIS   264 ...       |   Z      3.
16 1   HHH O      Zn    4    6   -1-413   -1   -1   -1 eee . 5 1bkc   0 0.4 2.0   ZN     1 HISE 405 ... .    |   U
17 1   HHHOS     Zn    4    6 -99    0   -1   -1   -1 eee.. 5 1kap   0 0.2 1.6   ZN   613 HISP 176 .....    |          3.
18 1   HHHOY     Zn    4    6 -99    0   -1   -1   -1 eee.. 5 1atl   0 0.1 1.8   ZN   401 HISA 142 .....    |          3.
```

```
all 18 same  over 0-4 -  alpha helix
   relseq      meanfi            meanpsi       number
       0         -62(  5)         -48(  6)          18         a
       1         -61(  6)         -38(  7)          18         a
       2         -62(  3)         -32( 10)          18         a
       3         -60(  6)         -33( 14)          18         a
       4         -65( 10)         -26( 42)          18         a
model : 1c7k at   83 A   1.00   (7)
```

1rmd is worst fit, then 1ali

note quite a few k conformations at relseq 3 and 4

local conformations nearly all quite different 40-90 deg rms.

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

info from 'old sequences files'   (with different nunbering of protein chains,
but almost exactly same ordering apart from 1rmd/1ali

SCOP fold info

==============

1,2,4 1ali:Alkaline phosphatase-like   core:3 layers: a/b/a; mixed beta-sheet of 8 strands, order 43516728, strand 7 is antipar   allel
to the rest

3 1rmd: 1-86 RING finger domain, C3HC4   dimetal(zinc)-bound alpha+beta motif; structurally diverse

```
          87-116 C2H2 and C2HC zinc fingers         alpha+beta metal(zinc)-bound fold: beta-hairpin + alpha-helix

5 1bs4:  Peptide deformylase                alpha-beta(5)-alpha; 3 layers: a/b/a; meander beta-sheet wraps around the C-terminal alp                              ha-hel

ix

8 1j98:  LuxS/MPP-like metallohydrolase core: beta-alpha-beta(2)-alpha(2); 2 layers: alpha/beta

9 1c7k:  Zincin-like                          contains mixed beta sheet with connection over free side of the sheet

10 1g12: Zincin-like                          contains mixed beta sheet with connection over free side of the sheet

11 1hs6: Zincin like     [ one of three entries, this one includes right  aa seqno]

12 1ezm: Zincin-like                          contains mixed beta sheet with connection over free side of the sheet

13 1i1i: Zincin-like                          contains mixed beta sheet with connection over free side of the sheet

14 1dmt: Zincin-like                          contains mixed beta sheet with connection over free side of the sheet

15 1hr6: LuxS/MPP-like metallohydrolase core: beta-alpha-beta(2)-alpha(2); 2 layers: alpha/beta

19 1i76: Zincin-like                          contains mixed beta sheet with connection over free side of the sheet

20 1lml: Zincin-like                          contains mixed beta sheet with connection over free side of the sheet

21 1bkc: Zincin-like                          contains mixed beta sheet with connection over free side of the sheet

25 1kap: Zincin-like                          [choose n terminal domain]

26 1atl: Zincin-like


Scop would put 8 15 in a group

                    9 10 11 12 13 14 19 20 21 25 26 in a group

      i.e. eleven are zincin like, rest various


Cath classification

===================                                       CATH code

1,2,4 1ali  COMPLEX (ZINC FINGER/DNA)             3.40.720.10      (446)

3      1rmd  DNA-BINDING PROTEIN                      3.30.40.10 (86) ,3.30.160.60 (30),

                                                                             5.1.2887.1 (116)

5      1bs4  HYDROLASE                                   3.90.45.10       (three domains   168 each)

8      1j98  SIGNALING PROTEIN  1.2 A               8.1.1.1            (154)

9      1c7k  HYDROLASE                                   3.40.390.10      (132)

10     1g12  HYDROLASE                                   6.1.211.10       (167)

11     1hs6  HYDROLASE                                   8.1.30.1          (610)

12     1ezm  HYDROLASE                                   1.10.390.10 (146), 3.10.170.10 (152),

                                                                            5.1.790.1  (290)
```

13	1ili	HYDROLASE	7.1.155.10	(665)
14	1dmt	HYDROLASE	7.1.25.10	(696)
15	1hr6	HYDROLASE	7.1.83.10	(457)
19	1i76	HYDROLASE	3.40.390.10	(163)
20	1lml	LEISHMANOLYSIN	2.10.55.10 (63), 2.30.34.10 (99), 3.10.170.20 (170) 3.90.132.10 (124), 5.1.2855.1 (465)	
21	1bkc	ZN-ENDOPEPTIDASE	3.40.390.10	(254,239,254,249)
25	1kap	ZINC METALLOPROTEASE	2.150.10.10 (237), 3.40.390.10 (233), 5.1.1240.1 (170)	
26	1atl	METALLOENDOPEPTIDASE	3.40.390.10	(200)

agreement of conformations
==========================

In 1ali there are three different HH4 groups within one chain - keep all

phi/psi agreement - rms difs

before i.e. 0-10 1ali all same (8,8,11 degs), otherwise different (80-100)

just before - from -3 1hr6 wildly different, others same (<=10 degs)

loop all same within 20 deg, many within 10 deg,

after to 14 all very different, except three within 1ali same

chi1 at 0 all 161 to 184

chi2 at 0 five -158 to -124 (1j98, 1hr6x4), twenty-two 77 to 127

chi1 at 4 all -52 to -110

chi2 at 4 five -134 to -127(1j98, 1hr6x4), twenty-one -62 to -38, one at -82 (1rmd)

coordination number, shape, distortion, N-M-N angle, comments on bond lengths etc
===

		CN	delta	his-M-his	unusual dists or ?
1	1ali	4	tet 5	100	
2		4	tet 4	109	
4		4	tet 5	100	
3	1rmd	4	tet 7	114	
5	1bs4	4	tet 9,9,9	104,104,106	cngroup includes O of ZN-O-ZN’ ?
8	1j98	4	tet 5	103	
9	1c7k*	4	tet 8	107	
10	1g12	6	oct 17	104	cngp includes bidentate asp
11	1hs6	5	tetp 9	103	+ long bond (2.8) to second O of glu
12	1ezm	4	tet 6	113	
13	1i1i	5	tetp 16	108	cngp includes bidentate glu
14	1dmt	5	tbp 20	101	cngp includes bidentate glu
15	1hr6	5	tetp 16,17,17,15	86,87,88,90	cngp includes bidentate glu
19	1i76	5	tetp 14	105	
20	1lml	4	tet 21	95	dif 0.27-0.38, shape odd, ? extra sym
21	1bkc	5	tetp 6,9,8,7	102,99,100,100	difs 0.26-0.39, addl donor at 2.9
25	1kap	5	tetp 14	104	
26	1atl	5	tetp 13,13	101,100	

* selected as model

old classification of conformation types

```
========================================
   start of selected part of cngroup          |          |
                             -10          -5           0           5          10          15
 1   1a1i from   143 A:   b b a b a a a a a a a a a a a a g b b a b b b a a
 2   1a1i from   171 A:   b b a b a a a a a a a a a a a b b b b b b b b b b
 3   1rmd from    98  :    b b b b a a a a a a a a a a b b b a b b b b b b b
 4   1a1i from   115 A:   b b a b a a a a a a a a a a a a g b b a b b b a a
 5   1bs4 from 1122 C:   b b j a a a a a a a a a a a a a a g b b a a a a b
 8   1j98 from    44 A:   g a b b b b a a a a a a a a a a a a a a a a a a a
```

```
 9  1c7k from   73 A:  a a b b a a a a a a a a a a a a a g b b b b a a b

10  1g12 from  107 A:  j b g b a a a a a a a a a a a b a a a g g b b b a

11  1hs6 from  285 A:  g a g a a a a a a a a a a a a a a b g a a b b b a

12  1ezm from  130  :  b b a b b a a a a a a a a a a a a a a a b a g b b

13  1i1i from  464 P:  b b a a a a a a a a a a a a a a a a a a a b b b a

14  1dmt from  573 A:  a a a a a a a a a a a a a a a a a b a a a a a b b

15  1hr6 from   60 B:  a b b a a a b g a a a a a a a a a a a j b a a b b

19  1i76 from  187 A:  b b b b a a a a a a a a a a a a a g b b b b a b a

20  1lml from  254  :  a b a a a a a a a a a a a a a a a g b b a a a a a

21  1bkc from  395 E:  b a a a a a a a a a a a a a a a a g b b b b b b g

25  1kap from  166 P:  b g b a a a a a a a a a a a a a a g b b b a b b b

26  1atl from  132 A:  b a a a a a a a a a a a a a a a a g b b b b b a a
```

```
Zn C C 2

  amino-acid sequences ****

   start of selected part of cngroup     |

   relseq             -10         -5          0          5         10         15

 1  1hw7 from  222 A:  V Y D P Q D V E F K C T C _ _ _ _ _ _ _ _ _

 2  1a1r from  113 A:  A P Q G S R S L T P C T C G S S D L Y L V T

 3  4mt2 from    3  :   P N C S C A T D G S C S C A G S C K C K Q C

 4  1pud from  308   :  S E D L K P L D S E C H C A V C Q K W S R A

 5  2hrv from   42 A:  N T V G D D Y I P S C D C T Q A T Y Y C K H

 6  1h7n from  123 A:  F P E L Y I I C D V C L C E Y T S H G H C G

 7  1i3j from  141 A:  _ _ _ _ _ _ _ _ K F C K C G V R I Q T S A Y

 8  1ile from  492   :  R P Y V D Q V E L A C A C G G T M R R V P Y

 9  2occ from   50 F:  P S I T N K R I V G C I C E E D N S T V I W


  aa types - acidic, etc **

   relseq             -10         -5          0          5         10         15

 1  1hw7 from  222 A:  h h a h p a h a h b h s h _ _ _ _ _ _ _ _ _

 2  1a1r from  113 A:  s h p g s b s h s h h s h g s s a h h h h s

 3  4mt2 from    3  :   h p h s h s s a g s h s h s g s h b h b p h

 4  1pud from  308   :  s a a h b h h a s a h p h s h h p b h s b s

 5  2hrv from   42 A:  p s h g a a h h h s h a h s p s s h h h b p

 6  1h7n from  123 A:  h h a h h h h h a h h h h a h s s p g p h g

 7  1i3j from  141 A:  _ _ _ _ _ _ _ _ b h h b h g h b h p s s s h

 8  1ile from  492   :  b h h h a p h a h s h s h g g s h b b h h h

 9  2occ from   50 F:  h s h s p b b h h g h h h a a a p s s h h h


  conformation sequences **

   relseq             -10         -5          0          5         10         15

 1  1hw7 from  222 A:  b b b b b b b b b b b k . . . . . . . . . .

 2  1a1r from  113 A:  b b b g b b b b b b b k k g b k b b b b b b

 3  4mt2 from    3  :   a k b k b b d k g k b d a g g k b k b a d b

 4  1pud from  308   :  a k b k b b a b k k b k b a a k k a b b a a
```

```
 5  2hrv from   42 A:  k b b b k b b b b d b k b k b b b b b a a
 6  1h7n from  123 A:  d k k b b b b b b b . k a k k b a k g b k b
 7  1i3j from  141 A:  . . . . . . . . . b b a k g b b b b k k b a
 8  1ile from  492  :  a a a a k k b b b b b a k . b b b b b b a b
 9  2occ from   50 F:  b b k . k b b b b b b b a b a d b k b b b b


 **********  protein names (pdb header) *******
 1  1hw7 at  232 A   2.2       CHAPERONE
 2  1a1r at  123 A   2.5       VIRAL PROTEIN
 3  4mt2 at   13      2.0       METALLOTHIONEIN
 4  1pud at  318      1.85      TRANSFERASE
 5  2hrv at   52 A   1.95      HYDROLASE
 6  1h7n at  133 A   1.6       DEHYDRATASE
 7  1i3j at  151 A   2.2       HYDROLASE/DNA
 8  1ile at  502      2.5       AMINOACYL-TRNA SYNTHETASE
 9  2occ at   60 F   2.3       OXIDOREDUCTASE


 **********  cngps  ************
 1 1  CC        Zn   2  -1  -1  -1  -1  -1  -1 ..     2 1hw7  0 0.2 2.2  ZN  240 CYSA 232 ..       |
 2 1  CCC       Zn   2  46  -1  -1  -1  -1  -1 ...   4 1a1r  0 0.2 2.5  ZN  901 CYSA 123 ...      |  Z
 3 2  CCCC      Zn   6   2  11  -1  -1  -1  -1 ....  4 4mt2  0 0.1 2.0  ZN   68 CYS    7 ....     |
 4 1  CCCH      Zn   2   3  26  -1  -1  -1  -1 ...d  4 1pud  0 0.1 1.9  ZN  400 CYS  318 ....     |        2.
 5 1  CCCH      Zn   2  58   2  -1  -1  -1  -1 ...d  4 2hrv  0 0.1 2.0  ZN  143 CYSA  52 ....     |        3.
 6 1  CCC       Zn   2   8  -1  -1  -1  -1  -1 ...   4 1h7n  0 0.1 1.6  ZN 1342 CYSA 133 ...      |  Z     4.
 7 1  CCCC      Zn   2  11   3  -1  -1  -1  -1 ....  4 1i3j  0 0.1 2.2  ZN  100 CYSA 151 ....     |        3.
 8 3  CCCC      Zn   3  38   2  -1  -1  -1  -1 ....  5 1ile  0 0.8 2.5  ZN 1102 CYS  461 ....     |   Z
 9 1  CCCC      Zn   2  20   3  -1  -1  -1  -1 ....  4 2occ  0 0.1 2.3  ZN   99 CYSF  60 ....     |        1.
```

over loop (phi 1-2, psi 0-1)

prot chains 1 and 2 agree within 11 deg

1-6 agree within 30-35 deg

7,8 agree with each other withion 26 deg

9 is slightly similar to 3 (31 deg)

```
for p chains 1-6

   relseq      meanfi             meanpsi        number

        0         -93( 38)         157( 21)           6          b

        1       -120( 16)           17( 15)           6          k

        2         -69( 22)           48( 98)           5          d


for p chains 1 and 2

   relseq      meanfi             meanpsi        number

        0         -72(  4)         153(  4)           2          b

        1       -138(  3)           19( 12)           2          k

        2       -104(  3)           -3( 12)           1          k


for p chains 7,8

 relseq      meanfi             meanpsi        number

        0         -69( 13)         136( 34)           2          b

        1         -50(  5)         -30(  2)           2          a

        2         -62( 14)         -16(  3)           2          k


no similarity in local conformations in any of these groups
```

```
Zn C C 3


  amino-acid sequences ****
   start of selected part of cngroup       |
   relseq             -10         -5          0          5          10          15
 1  1vfy from  182 A:  K F S L L N R K H H C R S C G G V F C Q E H S
 2  1f4l from  135 A:  F L P D R F V K G T C P K C K S P D Q Y G D N
 3  1zin from  120  :  L M E R L T G R R I C R N C G A T Y H L I F H
 4  1hxr from   84 A:  K D V G N V K F L V C A D C E I G P I G W H C
 5  1qf8 from  127 A:  I P G E A _ V K L Y C P K C _ D V Y T P K S S
 6  1vfy from  166 A:  _ _ _ D W I D S D A C M I C S K K F S L L N R
 7  1ptq from  234  :  K V Y N Y M S P T F C D H C G S L L W G L V K
 8  1a1i from  127 A:  R I H T G Q K P F Q C R I C M R N F S R S D H
 9  2occ from   72 F:  F W L H K G E A Q R C P S C G T H Y K L V P H
10  1hc7 from  448 A:  P F E A E P E E G F C V R C G R P S A Y G K R
11  1rmd from   51  :  L R C L K V M G S Y C P S C R Y P C F P T D L
12  1j8f from  211 C:  E K I F S E V T P K C E D C Q S L V K P D I V
13  1ptq from  251  :  L W G L V K Q G L K C E D C G M N V H H K C R
14  1dsz from 1125 A:  _ _ _ _ _ _ _ _ _ P C F V C Q D K S S G Y H Y
15  1g73 from  290 C:  Y A L G E G D K V K C F H C G G G L T D W K P
16  1rmd from   16  :  F P A H F V K S I S C Q I C E H I L A D P V E
17  1e7l from   13 A:  Q K F Y D A Q N G K C L I C Q R E L N P D V Q
18  1d09 from  128 B:  R K R A N D I A L K C K Y C E K E F S H N V V
19  1zin from  140  :  I F H P P A K P G V C D K C G G E L Y Q R A D
20  1f4l from  148 A:  C K S P D Q Y G D N C E V C G A T Y S P T E L
21  1ee8 from  228 A:  H A V Y G R E G L P C P A C G R P V E R R V V
22  1hxr from   13 A:  A E G R N R K A V L C Q R C G S R V L Q P G T
23  1j8f from  211 A:  E K I F S E V T P K C E D C Q S L V K P D I V
24  1d0q from   51 A:  S V S P E K Q I F H C F G C G A G G N A F T F
25  1fn9 from   41 A:  D M M V C G G A V V C M H C L G V V G S L Q R
26  1dsz from 1177 A:  C I I N K V T R N R C Q Y C R L Q K C F E V G
27  1gpc from   77  :  C S S T H G D Y D S C P V C Q Y I S K N D L Y
```

```
28  1ee8 from  248 A:  R V V A G R G T H F C P T C Q G E G P _ _ _ _
29  1guq from   42 A:  K Q V L P A H D P D C F L C A G N V R V T G D
30  1dcq from  254 A:  E V Q R M T G N D V C C D C G A P D P T W L S
31  1vfy from  212 A:  D L G I Y E P V R V C D S C F E D Y E F I V T
32  1a5t from   52  :  Q P Q G H K S C G H C R G C Q L M Q A G T H P
33  1zme from   40 C:  R H I K C P G G N P C Q K C V T S N A I C E Y
34  1g71 from  104 A:  K R C N H E P G T V C P I C L E D A K E L A K
35  1ctt from  119  :  L A A I T V N Y T P C G H C R Q F M N E L N S
36  1dsz from 1142 A:  S S G Y H Y G V S A C E G C K G F F R R S I Q
37  1dcq from  274 A:  W L S T N L G I L T C I E C S G I H R E L G V
38  1e7l from   48 A:  P K A G K V R G L L C N L C D A A E G Q M K H
39  1pud from  310  :  D L K P L D S E C H C A V C Q K W S R A Y I H
40  1hwt from   54 C:  _ _ _ _ _ R I P L S C T I C R K R K V K C D K
41  1hwt from   71 C:  K V K C D K L R P H C Q Q C T K T G V A H L C
42  1e3j from   89 A:  A V E P G V P C R R C Q F C K E G K Y N L C P
43  1het from   90 A:  I P L F T P Q C G K C R V C K H P E G N F C L
44  1i3j from  154 A:  G V R I Q T S A Y T C S K C R N R S G E N N S
45  1zme from   24 C:  _ _ _ _ _ _ _ S V A C L S C R K R H I K C P G
46  1rmd from   36  :  P V E T S C K H L F C R I C I L R C L K V M G
47  1a8h from  117  :  I Y Y G E Y E G L Y C V S C E R F Y T E K E L
48  1e3j from   86 A:  D R V A V E P G V P C R R C Q F C K E G K Y N
49  1het from   87 A:  D K V I P L F T P Q C G K C R V C K H P E G N
50  1a5t from   49  :  L C Q Q P Q G H K S C G H C R G C Q L M Q A G


  aa types - acidic, etc **
   relseq             -10         -5          0           5        10         15
 1  1vfy from  182 A:  b h s h h p b b p p h b s h g g h h h p a p s
 2  1f4l from  135 A:  h h h a b h h b g s h h b h b s h a p h g a p
 3  1zin from  120  :  h h a b h s g b b h h b p h g s s h p h h h p
 4  1hxr from   84 A:  b a h g p h b h h h h s a h a h g h h g h p h
 5  1qf8 from  127 A:  h h g a s _ h b h h h h b h _ a h h s h b s s
 6  1vfy from  166 A:  _ _ _ a h h a s a s h h h h s b b h s h h p b
 7  1ptq from  234  :  b h h p h h s h s h h a p h g s h h h g h h b
```

```
 8  1a1i from  127 A:  b h p s g p b h h p h b h h h b p h s b s a p

 9  2occ from   72 F:  h h h p b g a s p b h h s h g s p h b h h h p

10  1hc7 from  448 A:  h h a s a h a a g h h h b h g b h s s h g b b

11  1rmd from   51  :  h b h h b h h g s h h h s h b h h h h h s a h

12  1j8f from  211 C:  a b h h s a h s h b h a a h p s h h b h a h h

13  1ptq from  251  :  h h g h h b p g h b h a a h g h p h p p b h b

14  1dsz from 1125 A:  _ _ _ _ _ _ _ _ _ h h h h h p a b s s g h p h

15  1g73 from  290 C:  h s h g a g a b h b h h p h g g g h s a h b h

16  1rmd from   16  :  h h s p h h b s h s h p h h a p h h s a h h a

17  1e7l from   13 A:  p b h h a s p p g b h h h h p b a h p h a h p

18  1d09 from  128 B:  b b b s p a h s h b h b h h a b a h s p p h h

19  1zin from  140  :  h h p h h s b h g h h a b h g g a h h p b s a

20  1f4l from  148 A:  h b s h a p h g a p h a h h g s s h s h s a h

21  1ee8 from  228 A:  p s h h g b a g h h h h s h g b h h a b b h h

22  1hxr from   13 A:  s a g b p b b s h h h p b h g s b h h p h g s

23  1j8f from  211 A:  a b h h s a h s h b h a a h p s h h b h a h h

24  1d0q from   51 A:  s h s h a b p h h p h h g h g s g g p s h s h

25  1fn9 from   41 A:  a h h h h g g s h h h h p h h g h h g s h p b

26  1dsz from 1177 A:  h h h p b h s b p b h p h h b h p b h h a h g

27  1gpc from   77  :  h s s s p g a h a s h h h h p h h s b p a h h

28  1ee8 from  248 A:  b h h s g b g s p h h h s h p g a g h _ _ _ _

29  1guq from   42 A:  b p h h h s p a h a h h h h s g p h b h s g a

30  1dcq from  254 A:  a h p b h s g p a h h h a h g s h a h s h h s

31  1vfy from  212 A:  a h g h h a h h b h h a s h h a a h a h h h s

32  1a5t from   52  :  p h p g p b s h g p h b g h p h h p s g s p h

33  1zme from   40 C:  b p h b h h g g p h h p b h h s s p s h h a h

34  1g71 from  104 A:  b b h p p a h g s h h h h h h a a s b a h s b

35  1ctt from  119  :  h s s h s h p h s h h g p h b p h h p a h p s

36  1dsz from 1142 A:  s s g h p h g h s s h a g h b g h h b b s h p

37  1dcq from  274 A:  h h s s p h g h h s h h a h s g h p b a h g h

38  1e7l from   48 A:  h b s g b h b g h h h p h h a s s a g p h b p

39  1pud from  310  :  a h b h h a s a h p h s h h p b h s b s h h p

40  1hwt from   54 C:  _ _ _ _ _ b h h h s h s h h b b b b h b h a b
```

```
41  1hwt from   71 C:  b h b h a b h b h p h p p h s b s g h s p h h
42  1e3j from   89 A:  s h a h g h h h b b h p h h b a g b h p h h h
43  1het from   90 A:  h h h h s h p h g b h b h h b p h a g p h h h
44  1i3j from  154 A:  g h b h p s s s h s h s b h b p b s g a p p s
45  1zme from   24 C:  _ _ _ _ _ _ _ s h s h h s h b b b p h b h h g
46  1rmd from   36  :  h h a s s h b p h h h b h h h h b h h b h h g
47  1a8h from  117  :  h h h g a h a g h h h h s h a b h h s a b a h
48  1e3j from   86 A:  a b h s h a h g h h h b b h p h h b a g b h p
49  1het from   87 A:  a b h h h h h s h p h g b h b h h b p h a g p
50  1a5t from   49  :  h h p p h p g p b s h g p h b g h p h h p s g


   conformation sequences **
    relseq              -10       -5        0         5         10        15
 1   1vfy from  182 A:  b b b a k b b b b b b k a k g b b b b a a k d
 2   1f4l from  135 A:  b b b a k k b b . b b a a k g b b g b b j k b
 3   1zin from  120  :  a a a a a a k b b b b a a k g b b b b a a a .
 4   1hxr from   84 A:  b b b g g b b b b b b a a k g b . b k . b b b
 5   1qf8 from  127 A:  b b g a . . . b b b b a a . . . b b b b a b k
 6   1vfy from  166 A:  . . . . b b b . k b b k a k g b b b b a k b b
 7   1ptq from  234  :  b b b b b k b b b b b k a k g b b b b j a b g
 8   1a1i from  127 A:  a a a k g b b k b b b k a k g b b b a b a a a
 9   2occ from   72 F:  . b b b b . b b b b b a a k g b b b b b b b b
10  1hc7 from  448 A:  b a b . b b b b . b b a a k g b b b b g g . b
11  1rmd from   51  :  a a a a a a a j d b b a a k g b b b b a k k b
12  1j8f from  211 C:  a a a k k g b b b b b k a k g b b b b b d b b
13  1ptq from  251  :  b b j a b g a . b b b a a k g b b b b a k k a
14  1dsz from 1125 A:  . . . . . . . . . . b a a k g b b b k j b b g
15  1g73 from  290 C:  b b k g b j k b b b b a a k g b b b b g b b a
16  1rmd from   16  :  b b a a a a a a k b b k a k g b b b b g b b b
17  1e7l from   13 A:  a a a a a a k g g b b a a k g b b b b k b a k
18  1d09 from  128 B:  b b a . a b b b b b b a a k g b b b b k a a b
19  1zin from  140  :  a a . b b a b b g b b a a k g b b b b b b a k
20  1f4l from  148 A:  k g b b g b b j k b b k a k g b b . b k k k b
```

```
21  1ee8 from  228 A:  b a . b g k b g b b b k a k g b b b b b b b b

22  1hxr from   13 A:  a k g b b a b b b b b k a k g b b a b b b g k

23  1j8f from  211 A:  a a a a k g b b b b b k a k g b b b b b d b b

24  1d0q from   51 A:  b b b a a k g b b . b k k . g b . j b a a a a

25  1fn9 from   41 A:  a b b b b . k b b b b k k k g b b k . b a a k

26  1dsz from 1177 A:  b b b b a k k a k a b a a a a a a a a a a k g

27  1gpc from   77  :  b k a k d g d a a k b a a a a a a a a k g a a

28  1ee8 from  248 A:  b b b g g . b b b . b a a a d g . . . . . . .

29  1guq from   42 A:  b a b b b k b b a a b k k b b g b b b a k g b

30  1dcq from  254 A:  a a a k b b g k k b b a k k g b b g b k b b b

31  1vfy from  212 A:  k k g b k b b b b b b a a a a a a a a a a a k

32  1a5t from   52  :  d b b . k b b a g a b a a a a a a a a g k b k

33  1zme from   40 C:  k g b b b k j j . b b a a a a a k g b b b b b

34  1g71 from  104 A:  b a b d b b b g b b b a a a a a a a a a a a a

35  1ctt from  119  :  b a . b b b k b b b b a a a a a a a k a b k k

36  1dsz from 1142 A:  b k j b b g g b b b b a a a a a a a a a a a a

37  1dcq from  274 A:  b b b a k k g b b b b a a a a a a a a a a j a

38  1e7l from   48 A:  a k b g b b a . b b b a a a a a a a a a a a a

39  1pud from  310  :  b k b b a b k k b k b a a k k a b b a a a a a

40  1hwt from   54 C:  . . . . . . b b k b b a a a a a a g b b b k g

41  1hwt from   71 C:  g b b b k g b b a b b a a a a a k g k a k k b

42  1e3j from   89 A:  b b b k b b b a g k b a a a a k g d a k k b a

43  1het from   90 A:  b b b a b b b a g k b a a a k b a a b k d b k

44  1i3j from  154 A:  g b b b b k k b a b b a a k b b k b . k k b k

45  1zme from   24 C:  . . . . . . . . b b b a a a a a k g b b b k j

46  1rmd from   36  :  b b b b a k g b b b b a a a a a a a a a a a j

47  1a8h from  117  :  b b b b b b b . b b b a a a g b b b b a a a b

48  1e3j from   86 A:  b b b b b b k b b b a g k b a a a a k g d a k

49  1het from   87 A:  b b b b b b a b b b a g k b a a a k b a a b k

50  1a5t from   49  :  k b k d b b . k b b a g a b a a a a a a a a g


  **********  protein names (pdb header) *******

  1  1vfy at  192 A   1.15     TRANSPORT PROTEIN
```

```
  2  1f4l at  145 A   1.85     HYDROLASE
  3  1zin at  130     1.6      PHOSPHOTRANSFERASE
  4  1hxr at   94 A   1.65     METAL BINDING PROTEIN
  5  1qf8 at  137 A   1.74     TRANSFERASE
  6  1vfy at  176 A   1.15     TRANSPORT PROTEIN
  7  1ptq at  244     1.95     PHOSPHOTRANSFERASE
  8  1a1i at  137 A   1.6      COMPLEX (ZINC FINGER/DNA)
  9  2occ at   82 F   2.3      OXIDOREDUCTASE
 10  1hc7 at  458 A   2.43     AMINOACYL-TRNA SYNTHETASE
 11  1rmd at   61     2.1      DNA-BINDING PROTEIN
 12  1j8f at  221 C   1.70     GENE REGULATION, TRANSFERASE
 13  1ptq at  261     1.95     PHOSPHOTRANSFERASE
 14  1dsz at 1135 A   1.70     TRANSCRIPTION/DNA
 15  1g73 at  300 C   2.00     APOPTOSIS/APOPTOSIS INHIBITOR
 16  1rmd at   26     2.1      DNA-BINDING PROTEIN
 17  1e7l at   23 A   1.32     ENDONUCLEASE
 18  1d09 at  138 B   2.10     TRANSFERASE
 19  1zin at  150     1.6      PHOSPHOTRANSFERASE
 20  1f4l at  158 A   1.85     HYDROLASE
 21  1ee8 at  238 A   1.90     DNA BINDING PROTEIN
 22  1hxr at   23 A   1.65     METAL BINDING PROTEIN
 23  1j8f at  221 A   1.70     GENE REGULATION, TRANSFERASE
 24  1d0q at   61 A   1.71     TRANSFERASE
 25  1fn9 at   51 A   1.80     VIRUS/VIRAL PROTEIN
 26  1dsz at 1187 A   1.70     TRANSCRIPTION/DNA
 27  1gpc at   87     2.2      DNA BINDING PROTEIN
 28  1ee8 at  258 A   1.90     DNA BINDING PROTEIN
 29  1guq at   52 A   1.8      NUCLEOTIDYLTRANSFERASE
 30  1dcq at  264 A   2.10     METAL BINDING PROTEIN
 31  1vfy at  222 A   1.15     TRANSPORT PROTEIN
 32  1a5t at   62     2.2      ZINC FINGER
 33  1zme at   50 C   2.5      COMPLEX (TRANSCRIPTION REGULATION/DN
 34  1g71 at  114 A   2.30     REPLICATION
```

```
35  1ctt  at   129      2.2       HYDROLASE

36  1dsz  at  1152  A    1.70      TRANSCRIPTION/DNA

37  1dcq  at   284  A    2.10      METAL BINDING PROTEIN

38  1e7l  at    58  A    1.32      ENDONUCLEASE

39  1pud  at   320       1.85      TRANSFERASE

40  1hwt  at    64  C    2.5       COMPLEX (ACTIVATOR/DNA)

41  1hwt  at    81  C    2.5       COMPLEX (ACTIVATOR/DNA)

42  1e3j  at    99  A    2.3       OXIDOREDUCTASE

43  1het  at   100  A    1.15      OXIDOREDUCTASE

44  1i3j  at   164  A    2.2       HYDROLASE/DNA

45  1zme  at    34  C    2.5       COMPLEX (TRANSCRIPTION REGULATION/DN

46  1rmd  at    46       2.1       DNA-BINDING PROTEIN

47  1a8h  at   127       2.0       AMINOACYL-TRNA SYNTHETASE

48  1e3j  at    96  A    2.3       OXIDOREDUCTASE

49  1het  at    97  A    1.15      OXIDOREDUCTASE

50  1a5t  at    59       2.2       ZINC FINGER


 **********  cngps  ************

 1 1   CCCC      Zn    3  27    3  -1  -1  -1  -1 ....  4 1vfy  0 0.1 1.1  ZN  300 CYSA 192 ....     |

 2 1   CCCC      Zn    3  10    3  -1  -1  -1  -1 ....  4 1f4l  0 0.3 1.9  ZN  701 CYSA 145 ....     |         6.

 3 1   CCCC      Zn    3  17    3  -1  -1  -1  -1 ....  4 1zin  0 0.1 1.6  ZN  219 CYS  130 ....     |         2.

 4 3   CCCC      Zn    3  68    3  -1  -1  -1  -1 ....  4 1hxr  0 0.1 1.6  ZN  200 CYSA  23 ....     |

 5 3   CCCC      Zn    5  23    3  -1  -1  -1  -1 ....  4 1qf8  0 0.1 1.7  ZN  216 CYSA 109 ....     |         2.

 6 1   CCCH      Zn    3  21    3  -1  -1  -1  -1 ...d  4 1vfy  0 0.1 1.1  ZN  301 CYSA 176 ....     |

 7 1   CCHC      Zn    3  22    3  -1  -1  -1  -1 ..d.  4 1ptq  0 0.3 2.0  ZN    2 CYS  244 ....     |         2.

 8 1   CCHH      Zn    3  13    4  -1  -1  -1  -1 ..ee  4 1a1i  0 0.1 1.6  ZN  202 CYSA 137 ....     |

 9 3   CCCC      Zn    2  20    3  -1  -1  -1  -1 ....  4 2occ  0 0.1 2.3  ZN   99 CYSF  60 ....     |         1.

10 3   CCCC      Zn    5  26    3  -1  -1  -1  -1 ....  4 1hc7  0 0.1 2.4  ZN  490 CYSA 427 ....     |         6.

11 3   CHCC      Zn    2  18    3  -1  -1  -1  -1 .d..  4 1rmd  0 0.1 2.1  ZN  119 CYS   41 ....     |

12 3   CCCC      Zn    5  21    3  -1  -1  -1  -1 ....  4 1j8f  0 0.1 1.7  ZN 3001 CYSC 195 ....     |

13 2   HCCC      Zn   30    3  16  -1  -1  -1  -1 d...  4 1ptq  0 0.3 2.0  ZN    1 HIS  231 ....     |         2.

14 1   CCCC      Zn    3  14    3  -1  -1  -1  -1 ....  4 1dsz  0 0.1 1.7  ZN 1121 CYSA1135 ....     |

15 1   CCHC      Zn    3  17    7  -1  -1  -1  -1 ..e.  4 1g73  0 0.2 2.0  ZN  501 CYSC 300 ....     |
```

```
16 1  CCCC      Zn   3  17   3  -1  -1  -1  -1 ....  4 1rmd  0 0.1 2.1  ZN  118 CYS   26 ....    |
17 1  CCCC      Zn   3  32   3  -1  -1  -1  -1 ....  4 1e7l  0 0.1 1.3  ZN 1165 CYSA  23 ....    |        3.
18 3  CCCC      Zn   5  24   3  -1  -1  -1  -1 ....  4 1d09  0 0.1 2.1  ZN 1313 CYSB 109 ....    |        2.
19 3  CCCC      Zn   3  17   3  -1  -1  -1  -1 ....  4 1zin  0 0.1 1.6  ZN  219 CYS  130 ....    |        2.
20 3  CCCC      Zn   3  10   3  -1  -1  -1  -1 ....  4 1f4l  0 0.3 1.9  ZN  701 CYSA 145 ....    |        6.
21 1  CCCC      Zn   3  17   3  -1  -1  -1  -1 ....  4 1ee8  0 0.1 1.9  ZN  501 CYSA 238 ....    |
22 1  CCCC      Zn   3  68   3  -1  -1  -1  -1 ....  4 1hxr  0 0.1 1.6  ZN  200 CYSA  23 ....    |
23 3  CCCC      Zn   5  21   3  -1  -1  -1  -1 ....  4 1j8f  0 0.1 1.7  ZN 1001 CYSA 195 ....    |
24 3  CHCC      Zn   3  18   3  -1  -1  -1  -1 .d..  4 1d0q  0 0.2 1.7  ZN  151 CYSA  40 ....    |        2.
25 1  CCHC      Zn   3  17   2  -1  -1  -1  -1 ..e.  4 1fn9  0 0.2 1.8  ZN 1001 CYSA  51 ....    |
26 3  CCCC      Zn   6  10   3  -1  -1  -1  -1 ....  4 1dsz  0 0.1 1.7  ZN 1122 CYSA1171 ....    |
27 3  HCCC      Zn  13  10   3  -1  -1  -1  -1 e...  4 1gpc  0 0.4 2.2  ZN    1 HIS   64 ....    |
28 3  CCCC      Zn   3  17   3  -1  -1  -1  -1 ....  4 1ee8  0 0.1 1.9  ZN  501 CYSA 238 ....    |
29 1  CCHH      Zn   3  60  49  -1  -1  -1  -1 ..dd  4 1guq  0 0.1 1.8  ZN  350 CYSA  52 ....    |        2.
30 1  CCCC      Zn   3  17   3  -1  -1  -1  -1 ....  4 1dcq  0 0.1 2.1  ZN  600 CYSA 264 ....    |
31 3  CCCC      Zn   3  27   3  -1  -1  -1  -1 ....  4 1vfy  0 0.1 1.1  ZN  300 CYSA 192 ....    |
32 3  CCCC      Zn   9   3   3  -1  -1  -1  -1 ....  4 1a5t  0 0.1 2.2  ZN  501 CYS   50 ....    |        2.
33 2  CCCC      Zn  16   3   7  -1  -1  -1  -1 ....  5 1zme  0 0.4 2.5  ZN    2 CYSC  34 ....    |   J
34 3  CHCC      Zn   2   6   3  -1  -1  -1  -1 .e..  4 1g71  0 0.2 2.3  ZN  339 CYSA 106 ....    |
35 2  HCC       Zn  27   3  -1  -1  -1  -1  -1 d..   4 1ctt  0 0.2 2.2  ZN  296 HIS  102 ...     |  Z    3.
36 3  CCCC      Zn   3  14   3  -1  -1  -1  -1 ....  4 1dsz  0 0.1 1.7  ZN 1121 CYSA1135 ....    |
37 3  CCCC      Zn   3  17   3  -1  -1  -1  -1 ....  4 1dcq  0 0.1 2.1  ZN  600 CYSA 264 ....    |
38 3  CCCC      Zn   3  32   3  -1  -1  -1  -1 ....  4 1e7l  0 0.1 1.3  ZN 1165 CYSA  23 ....    |        3.
39 2  CCCH      Zn   2   3  26  -1  -1  -1  -1 ...d  4 1pud  0 0.1 1.9  ZN  400 CYS  318 ....    |        2.
40 1  CCCC      Zn   3   7   7  -1  -1  -1  -1 ....  5 1hwt  0 0.3 2.5  ZN    1 CYSC  64 ....    |   J
41 2  CCCC      Zn  17   3   9  -1  -1  -1  -1 ....  5 1hwt  0 0.3 2.5  ZN    2 CYSC  64 ....    |   J
42 2  CCCC      Zn   3   3   8  -1  -1  -1  -1 ....  4 1e3j  0 0.2 2.3  ZN  901 CYSA  96 ....    |
43 2  CCCC      Zn   3   3   8  -1  -1  -1  -1 ....  4 1het  0 0.2 1.1  ZN  401 CYSA  97 ....    |        1.
44 3  CCCC      Zn   2  11   3  -1  -1  -1  -1 ....  4 1i3j  0 0.1 2.2  ZN  100 CYSA 151 ....    |        3.
45 1  CCCC      Zn   3   7   6  -1  -1  -1  -1 ....  5 1zme  0 0.4 2.5  ZN    1 CYSC  34 ....    |   J
46 3  CCCC      Zn   3  17   3  -1  -1  -1  -1 ....  4 1rmd  0 0.1 2.1  ZN  118 CYS   26 ....    |
47 1  CCCH      Zn   3  14   3  -1  -1  -1  -1 ...d  4 1a8h  0 0.2 2.0  ZN 1000 CYS  127 ....    |        6.
48 1  CCCC      Zn   3   3   8  -1  -1  -1  -1 ....  4 1e3j  0 0.2 2.3  ZN  901 CYSA  96 ....    |
```

```
49 1  CCCC     Zn   3   3   8  -1  -1  -1  -1 ....  4 1het  0 0.2 1.1  ZN  401 CYSA  97 ....    |        1.
50 2  CCCC     Zn   9   3   3  -1  -1  -1  -1 ....  4 1a5t  0 0.1 2.2  ZN  501 CYS   50 ....    |        2.

3 cngps of 1ile removed, because of doubts about the structure determination accuracy



p chains 48,49 are very similar, and 50 close, over loop phi 1-3, psi 0-2
 relseq     meanfi           meanpsi      number
      0        -73(  4)        -27(  6)          3        a
      1          77(  7)          7(  7)          3        g
      2      -128(  8)        -21( 21)          3        a
      3        -73( 20)        171(  8)          3        b
    entirely different from others

p chains 1-47 are failry close over loop phi 1-3, psi 0-2
 relseq     meanfi           meanpsi      number
      0        -77( 23)        139( 19)         47        b
      1        -63(  9)        -25( 12)         47        a
      2        -80( 15)        -45( 13)         47        a
      3        -91( 26)        -13( 29)         46        k

 but within this there are subgroups
p chains 1-14
  relseq     meanfi           meanpsi      number
      0        -69( 14)        126(  4)         14        b
      1        -64(  6)        -20(  6)         14        a
      2        -88(  7)        -50(  6)         14        a
      3      -109( 10)          -4(  4)         13        k
    local confs similar (17 deg) for relseq-1 to 5, not outside this

p chains 31-46
  relseq     meanfi           meanpsi      number
      0        -76( 21)        163( 10)         16        b
```

```
      1          -58(   8)         -34(   5)          16         a

      2          -64(   5)         -45(   6)          16         a

      3          -68(   6)         -31( 11)          16         a

     no similarity in local conformations

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

```
Zn C C 5

   amino-acid sequences ****

    start of selected part of cngroup        |

    relseq              -10         -5          0          5         10         15

 1  4mt2 from   14  :  S C A G S C K C K Q C K C T S C K K S C C S C C P

 2  1j8f from  185 A:  E A H G T F Y T S H C V S A S C R H E Y P L S W M

 3  1qf8 from   99 A:  E K Y Q Q G D F G Y C P R V Y C E N Q P _ L P I G

 4  1a1i from   97 A:  _ _ _ _ _ _ R P Y A C P V E S C D R R F S R S A D

 5  1d09 from   99 B:  L P E R I D N V L V C P N S N C I S H A E P V S S

 6  4mt2 from    9  :  T D G S C S C A G S C K C K Q C K C T S C K K S C

 7  1hc7 from  417 A:  V Q E G F A L A F H C G D K A C E R L I Q E E T T

 8  1a73 from   90 A:  N G K T C T A S H L C H N T R C H N P L H L C W E

 9  1rmd from   81  :  L N I L N S L M V K C P A Q D C N E E V S L E K Y

10  1ed5 from   86 A:  L C A Q S Q Q D G P C T P R R C L G S L V L P R K


   aa types - acidic, etc **

    relseq              -10         -5          0          5         10         15

 1  4mt2 from   14  :  s h s g s h b h b p h b h s s h b b s h h s h h h

 2  1j8f from  185 A:  a s p g s h h s s p h h s s s h b p a h h h s h h

 3  1qf8 from   99 A:  a b h p p g a h g h h h b h h h a p p h _ h h h g

 4  1a1i from   97 A:  _ _ _ _ _ _ b h h s h h h a s h a b b h s b s s a

 5  1d09 from   99 B:  h h a b h a p h h h h h p s p h h s p s a h h s s

 6  4mt2 from    9  :  s a g s h s h s g s h b h b p h b h s s h b b s h

 7  1hc7 from  417 A:  h p a g h s h s h p h g a b s h a b h h p a a s s

 8  1a73 from   90 A:  p g b s h s s s p h h p p s b h p p h h p h h h a

 9  1rmd from   81  :  h p h h p s h h h b h h s p a h p a a h s h a b h

10  1ed5 from   86 A:  h h s p s p p a g h h s h b b h h g s h h h h b b


   conformation sequences **

    relseq              -10         -5          0          5         10         15

 1  4mt2 from   14  :  d a g g k b k b a d b k b k k b d b b k b k k b b

 2  1j8f from  185 A:  b a k j b k a b b b b k b a k a g b b b b a a a a
```

```
  3  1qf8 from    99 A:  a a a a k g a k j b b k b a k k g g b . . . b b b
  4  1a1i from    97 A:  . . . . . . . k b b b k b a k b k b b b a b a a a
  5  1d09 from    99 B:  b b k b b b g a b b b k b a k b a a a a b d b b b
  6  4mt2 from     9  :  d k g k b d a g g k b k b a d b k b k k b d b b k
  7  1hc7 from   417 A:  k a a . b b b b b b a g b a a a a a a a a a a k g
  8  1a73 from    90 A:  g g b b b b b b b k k g b a k b k b k k k b b b b
  9  1rmd from    81  :  a a a a k k b b b b b k b b g b d b b b b a a k a
 10  1ed5 from    86 A:  a a k k b k b b b k b b a a b b k a k b a d b d a

  **********  protein names (pdb header) *******
  1  4mt2 at    24       2.0        METALLOTHIONEIN
  2  1j8f at   195 A    1.70       GENE REGULATION, TRANSFERASE
  3  1qf8 at   109 A    1.74       TRANSFERASE
  4  1a1i at   107 A    1.6        COMPLEX (ZINC FINGER/DNA)
  5  1d09 at   109 B    2.10       TRANSFERASE
  6  4mt2 at    19       2.0        METALLOTHIONEIN
  7  1hc7 at   427 A    2.43       AMINOACYL-TRNA SYNTHETASE
  8  1a73 at   100 A    1.8        COMPLEX (HOMING ENDONUCLEASE/DNA)
  9  1rmd at    91       2.1        DNA-BINDING PROTEIN
 10  1ed5 at    96 A    1.80       OXIDOREDUCTASE

  **********  cngps  ************
  1 3   CCCC      Zn    4    5    5   -1   -1   -1   -1 ....   4 4mt2   0 0.2 2.0   ZN    67 CYS    15 ....     |
  2 1   CCCC      Zn    5   21    3   -1   -1   -1   -1 ....   4 1j8f   0 0.1 1.7   ZN 1001 CYSA 195 ....     |
  3 1   CCCC      Zn    5   23    3   -1   -1   -1   -1 ....   4 1qf8   0 0.1 1.7   ZN   216 CYSA 109 ....     |         2.
  4 1   CCHH      Zn    5   13    4   -1   -1   -1   -1 ..ee   4 1a1i   0 0.1 1.6   ZN   201 CYSA 107 ....     |
  5 1   CCCC      Zn    5   24    3   -1   -1   -1   -1 ....   4 1d09   0 0.1 2.1   ZN 1313 CYSB 109 ....     |         2.
  6 2   CCCC      Zn    4    5    5   -1   -1   -1   -1 ....   4 4mt2   0 0.2 2.0   ZN    67 CYS    15 ....     |
  7 1   CCCC      Zn    5   26    3   -1   -1   -1   -1 ....   4 1hc7   0 0.1 2.4   ZN   490 CYSA 427 ....     |         6.
  8 2   CCCH      Zn   59    5    5   -1   -1   -1   -1 ...d   4 1a73   0 0.1 1.8   ZN   202 CYSA  41 ....     |
  9 1   CCHH      Zn    5   12    4   -1   -1   -1   -1 ..ee   4 1rmd   0 0.1 2.1   ZN   120 CYS    91 ....     |
 10 1   CCCC      Zn    5 -99    5   -1   -1   -1   -1 ....   4 1ed5   0 0.1 1.8   ZN   900 CYSA  96 ....     |         1.
```

```
over loop (phi 1-5, psi 0-4)

p chains 1-4 are same within 17 deg

         1-5 are same within 20 deg


         7,8 are same within 22 deg but dif from above

         others all fairly different


for p chains 1-4   model 1a1i at  107 A   1.6

  relseq     meanfi          meanpsi      number

     0        -60(  4)       133(  5)        4       b

     1        -80(  8)        -8(  5)        4       k

     2        -70(  5)       127(  4)        4       b

     3        -56( 12)       -32( 18)        4       a

     4        -79(  5)        -4(  6)        4       k

     5       -107( 12)       135(118)        4       b

  local conformations all fairly different (> 60 deg)  and 5


for p chains 7 8

 relseq     meanfi          meanpsi      number

     0        -79( 22)       -15( 14)        2       k

     1          75( 28)        22( 14)        2       g

     2       -106( 17)       113( 20)        2       b

     3        -53(  4)       -45(  0)        2       a

     4        -62(  6)       -33( 19)        2       a

     5        -65(  4)        48(124)        2       d

   local confs not similar
```

```
Zn  H C 3


   amino-acid sequences ****
    start of selected part of cngroup         |
    relseq                -10          -5           0           5          10          15
 1  1ptq from  259  :  L K C E D C G M N V H H K C R E K V A N L C _
 2  1g5c from   77 A:  L G D N E I I I V G H T D C G M A R L D E D L
 3  1g5c from   77 C:  L G D N E I I I V G H T D C G M A R L D E D L
 4  1g5c from   77 E:  L G D N E I I I V G H T D C G M A R L D E D L
 5  1ekj from  210 D:  L K V S N I V V I G H S A C G G I K G L L S F


   aa types - acidic, etc **
    relseq                -10          -5           0           5          10          15
 1  1ptq from  259  :  h b h a a h g h p h p p b h b a b h s p h h _
 2  1g5c from   77 A:  h g a p a h h h h g p s a h g h s b h a a a h
 3  1g5c from   77 C:  h g a p a h h h h g p s a h g h s b h a a a h
 4  1g5c from   77 E:  h g a p a h h h h g p s a h g h s b h a a a h
 5  1ekj from  210 D:  h b h s p h h h h g p s s h g g h b g h h s h


   conformation sequences **
    relseq                -10          -5           0           5          10          15
 1  1ptq from  259  :  b b b a a k g b b b b a k k a k k b b a b . .
 2  1g5c from   77 A:  k g b k b b b b b b b b g a g k d . b d a a a
 3  1g5c from   77 C:  k g b k b b b b b b b b g b a k k d b b a a a
 4  1g5c from   77 E:  k g b k b b b b b b b b g b . b b a a b a a a
 5  1ekj from  210 D:  k g b k b b b b b b b b g b a a a a a a a a b


 **********  protein names (pdb header) *******
 1  1ptq at  269      1.95      PHOSPHOTRANSFERASE  Cysteine-rich domain
 2  1g5c at   87 A    2.10      LYASE
 3  1g5c at   87 C    2.10      LYASE
 4  1g5c at   87 E    2.10      LYASE         Resolvase-like
```

```
 5  1ekj at  220 D   1.93      LYASE         Resolvase-like


 **********  cngps  ************
 1 3  CCHC      Zn   3  22   3  -1  -1  -1  -1  -1 ..d.   4 1ptq  0 0.3 2.0  ZN    2 CYS  244 ....    |        2.
 2 2  CHC       Zn  55   3  -1  -1  -1  -1  -1 .e.    4 1g5c  0 0.1 2.1  ZN 1001 CYSA  32 ...     |  Z     4.
 3 2  CHC       Zn  55   3  -1  -1  -1  -1  -1 .e.    4 1g5c  0 0.1 2.1  ZN 1003 CYSC  32 ...     |  Z     4.
 4 2  CHC       Zn  55   3  -1  -1  -1  -1  -1 .e.    4 1g5c  0 0.1 2.1  ZN 1005 CYSE  32 ...     |  Z     4.
 5 2  CHC       Zn  60   3  -1  -1  -1  -1  -1 .e.    3 1ekj  0 0.1 1.9  ZN 4004 CYSD 160 ...     |        4.


Conformations at loop
 1 different, rest all same


for pc 2-5   model 1ekj at  220 D   1.93
 relseq     meanfi           meanpsi      number
      0      -150(  4)        156(  3)          4        b
      1        -67(  5)        153(  5)          4        b
      2          58(  2)         48(  5)          4        g
      3        -92( 19)        -95(166)          4        a


local conformations of 2-5 are fairly different  -  r.m.s. 35-60 deg
     but are same before loop



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

```
ZN   H H 2


   amino-acid sequences ****
    start of selected part of cngroup         |
    relseq               -10         -5          0          5         10         15
 1   1e4c from    82 P:  Q S R P D A N A V V H N H A V H C T A V S I
 2   1dq3 from   146 A:  I A G D G C F D K Y H S H V K G H E Y I Y D
 3   1j79 from     6 A:   V L K I R R P D D W H L H L R D G D M L K T
 4   1hzy from    45 B:  T I S E A G F T L T H E H I C G S S A G F L
 5   1bf6 from     2 A:   S F D P T G Y T L A H E H L H I D L S G F K
 6   1koq from    82 A:  N G R T Y T L K Q F H F H V P S E N Q I K G
 7   2bc2 from    76 A:  Q K R V T D V I I T H A H A D R I G G I K T
 8   1sml from    74 A:  P R D L R L I L L S H A H A D H A G P V A E
 9   1b66 from    38 B:  V F G K C N N P N G H G H N Y K V V V T I H
10   1ton from    87  :  Q S F R H P D Y P V H D H S N D L M L L H L
11   1qh5 from    44 B:  G V K L T T V L T T H H H W D H A G G N E K


   aa types - acidic, etc **
    relseq               -10         -5          0          5         10         15
 1   1e4c from    82 P:  p s b h a s p s h h p p p s h p h s s h s h
 2   1dq3 from   146 A:  h s g a g h h a b h p s p h b g p a h h h a
 3   1j79 from     6 A:   h h b h b b h a a h p h p h b a g a h h b s
 4   1hzy from    45 B:  s h s a s g h s h s p a p h h g s s s g h h
 5   1bf6 from     2 A:   s h a h s g h s h s p a p h p h a h s g h b
 6   1koq from    82 A:  p g b s h s h b p h p h p h h s a p p h b g
 7   2bc2 from    76 A:  p b b h s a h h h s p s p s a b h g g h b s
 8   1sml from    74 A:  h b a h b h h h h s p s p s a p s g h h s a
 9   1b66 from    38 B:  h h g b h p p h p g p g p p h b h h h s h p
10   1ton from    87  :  p s h b p h a h h h p a p s p a h h h h p h
11   1qh5 from    44 B:  g h b h s s h h s s p p p h a p s g g p a b


   conformation sequences **
```

```
   relseq              -10         -5          0          5         10         15

 1  1e4c from   82 P:  a a b a k b k b b b b b d b a a a a a a a a

 2  1dq3 from  146 A:  a a a k b b b . b b b a g b b g d b b b b b

 3  1j79 from    6 A:  b b b b b b b b b k . b b k j b . a a a a a

 4  1hzy from   45 B:  b a a k a j b b b k b b b a b b . b b g a a

 5  1bf6 from    2 A:  . b b k k b b b b k b b b a b b b k a a a a

 6  1koq from   82 A:  g g b b b b b a . b b b b b b b a b b b g g

 7  2bc2 from   76 A:  g b b b a b b b b a . d b a k a k a k a a a

 8  1sml from   74 A:  a a a b a b b b b a b d b a k a k a k a a a

 9  1b66 from   38 B:  a k . a k k b a k j . b b b b b b b b b b b

10  1ton from   87  :  . b b b b a k b . b b b k b g d b b b b b b

11  1qh5 from   44 B:  g b b b a b b b b a . k b a k a k k k a a a


 **********  protein names (pdb header) *******

 1  1e4c at   92 P   1.66     ALDOLASE (CLASS II)        Class II aldolase

 2  1dq3 at  156 A   2.10     HYDROLASE                     Homing endonuclease-like

 3  1j79 at   16 A   1.70     HYDROLASE                       TIM beta/alpha-barrel

 4  1hzy at   55 B   1.30     HYDROLASE                       TIM beta/alpha-barrel

 5  1bf6 at   12 A   1.7      PHOSPHOTRIESTERASE            TIM beta/alpha-barrel

 6  1koq at   92 A   1.90     LYASE                               Carbonic anhydrase

 7  2bc2 at   86 A   1.7      HYDROLASE                   Metallo-hydrolase/oxidoreductase

 8  1sml at   84 A   1.70     HYDROLASE                    Metallo-hydrolase/oxidoreductase

 9  1b66 at   48 B   1.9      TETRAHYDROBIOPTERIN BIOSYNTHESIS      T-fold

10  1ton at   97     1.8      HYDROLASE(SERINE PROTEINASE)   Trypsin-like serine proteases

11  1qh5 at   54 B   1.45     HYDROLASE                       Metallo-hydrolase/oxidoreductase


 **********  cngps  ************

 1 2  EHHH      Zn  19   2  61  -1  -1  -1  -1 .eee  5 1e4c  0 0.2 1.7  ZN  999 GLUP  73 ....    |   Z   4.

 2 1  HH         Zn   2  -1  -1  -1  -1  -1  -1 de     2 1dq3  0 0.2 2.1  ZN  901 HISA 156 ..       |

 3 1  HH D      Zn   2  -1 232  -1  -1  -1  -1 ee .  5 1j79  0 0.2 1.7  ZN  400 HISA  16 .. .     | U Z   3.

 4 1  HHD       Zn   2 244  -1  -1  -1  -1  -1 ee.   5 1hzy  0 0.1 1.3  ZN  401 HISB  55 ...      |  UZ   3.

 5 1  HHED      Zn   2 111 118  -1  -1  -1  -1 ee..  4 1bf6  0 0.3 1.7  ZN    1 HISA  12 ....    |

 6 1  HHH       Zn   2  17  -1  -1  -1  -1  -1 eed   4 1koq  0 0.2 1.9  ZN  301 HISA  92 ...      |  Z    4.
```

```
 7  1   HHH         Zn    2   61   -1   -1   -1   -1   -1  ede    5  2bc2   0  0.3  1.7   ZN     1  HISA   86  ...      |   ZZ    3.
 8  1   HHH         Zn    2   74   -1   -1   -1   -1   -1  ede    4  1sml   1  0.2  1.7   ZN   269  HISA   84  ...      |   Z     3.
 9  2   HHH         Zn   25    2   -1   -1   -1   -1   -1  eee    4  1b66   0  0.5  1.9   ZN   402  HISB   23  ...      |   U     4.
10  2   HHH         Zn   40    2   -1   -1   -1   -1   -1  eee    3  1ton   0  0.1  1.8   ZN   200  HIS    57  ...      |           N
11  1   HHHD        Zn    2   54   24   -1   -1   -1   -1  ede.   5  1qh5   0  0.3  1.5   ZN   261  HISB   54  ....     |   Z    3.


Over  loop
pchains 3 4 5 agree well, 1 and 10 a little less well with them
  model :   1hzy at    55 B    1.30   (4)
  relseq      meanfi           meanpsi       number
      0       -130( 40)        113( 20)          5         b
      1       -143(  6)        171( 13)          5         b
      2       -122( 19)         90( 56)          5         b
   but local conformations different    3 4 5 same fold



p chains   7 8 agree    bdb
p chains   6 9 agree    bbb   and are somewhat like 3 4 5

p chains 2 11 different from each other and above

------------------------------------------------------------------------------
info from old sequence file, some different numbering

loop region:  3,5,7,9,14    are   bbb      model 5   1hzy at 55B   (resoln 1.30)
   0-2          11,13,17      are   bab      model 17 1qh5 at 54B            1.45)
                 1,16                 bba             1   1e4c at 92P            1.66
                 2                     bag

  start of selected part of cngroup         |
                           -10          -5           0            5          10          15
 1  1e4c from    82 P:  a a b a a b a b b b b b a b a a a a a a a a
 2  1dq3 from   146 A:  a a a a b b b b b b b a g b b g a b b b b b
```

```
 3  1j79 from   6 A:  b b b b b b b b b a b b b a j b j a a a a a
 5  1hzy from  45 B:  b a a a a j b b b a b b b a b b b b b g a a
 7  1bf6 from   2 A:  b b b a a b b b b a b b b a b b b a a a a a
 9  1koq from  82 A:  g g b b b b b a b b b b b b b b a b b b g g
11  2bc2 from  76 A:  g b b b a b b b b a b a b a a a a a a a a a
13  1sml from  74 A:  a a a b a b b b b a b a b a a a a a a a a a
14  1b66 from  38 B:  a a j a a a b a a j b b b b b b b b b b b b
16  1ton from  87  :  b b b b b a a b j b b b a b g b b b b b b b
17  1qh5 from  44 B:  g b b b a b b b b a b a b a a a a a a a a a
```

		CN	delta	his-M-his	unusual dists or ?
1	1e4c	5	tetp 2.6	103	-
2	1dq3	2		123	??
3	1j79	5	tbp 6.7,7.7	115,121	-
5	1hzy	5	tbp 7,7.5	116,114	-
7	1bf6	4	sqp 25,25	119,120	looks like tbp, one don missing dists long too (.16-.2+)
9	1koq	4	tet 10,13	107,111	dif 0.1-0.3
11	2bc2	5	tbp 12,9	101,98	dif 0-0.27
13	1sml	4	tet 8	99	+ H2O at 3.04
14	1b66	5	tbp 13	114	.22-.30
		and 4	tet 11	115	.17-.22, missing donor ?
16	1ton	3		100	missing donor ?
17	1qh5	6	oct 16	95	.15-.22, or CN 5 plus H2O at 2.47
		and 5	tetp 11	95	.13-.23,

ARCHITECTURE OF METAL COORDINATION GROUPS - TABLE 5W

Proteins where there appears to be more than one metal coordination group associated with the protein.

Equivalent coordination groups involving duplicate protein chains within the asymmetric unit have been eliminated. The column headings have the same meaning as in Table 3 a) and Table 3D. Most cases are straightforward, with two or more metal ions coordinated by amino-acids of one protein chain. In a few cases (marked) there are two different protein chains, each with a metal site. In 11 cases (marked as ambiguities) the procedure has picked differences between nearly equivalent protein chains within the asymmetric unit. In a few of these the coordinating amino-acid is interpreted as asp in one, asn in another, or glu and gln; more often, an extra donor has been identified in one of the coordination groups.

Ccoordination groups where the metal is associated with two or more protein chains within the asymmetric unit are not included, but are listed separately.

```
cngroup       seqdif 1 to 7             his  cn     cn2 rms res  metal   startAA  carbi othdonors e.c.no

    1a2x
DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1a2x  2 0.2 2.3  CA  160 ASPA 103 ....b   |
DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1a2x  1 0.4 2.3  CA  161 ASPA 139 ..b.b   |
    1acc
DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1acc  1 0.2 2.1  CA  800 ASP  177 .....   |    Z
DDEOOD   Ca   2   7  34   3  10  -1  -1 ......6 1acc  0 0.3 2.1  CA  801 ASP  179 ..b...  |
    1alv
ODOE     Ca   3   2   5  -1  -1  -1  -1 ....  6 1alv  0 0.2 1.9  CA    1 ALAA 107 ...b    |   ZZ  3.4.22.17;
DDTOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1alv  0 0.2 1.9  CA    2 ASPA 150 ....b   |    Z  3.4.22.17;
DDSOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1alv  0 0.2 1.9  CA    3 ASPA 180 ....b   |    Z  3.4.22.17;
DDDN     Ca  88   2   1  -1  -1  -1  -1 ....  6 1alv  0 0.2 1.9  CA    4 ASPA 135 .bb.    |   ZZ  3.4.22.17;
    1aru
ODODS    Ca   0  18   2   2  -1  -1  -1 ..... 7 1aru  0 0.2 1.6  CA  346 ASP   57 .....   |    ZZ 1.11.1.7;
OSDOTOD  Ca   0  17   2   0   3   2  -1 ......7 1aru  0 0.2 1.6  CA  347 SER  185 ..b.... |       1.11.1.7;
    1ava
NDODO    Ca  47   3   7  35  -1  -1  -1 ..... 6 1ava  0 0.2 1.9  CA  500 ASNA  91 .b...   |    Z  3.2.1.1;
EOOD     Ca   3   2   4  -1  -1  -1  -1 ....  5 1ava  1 0.2 1.9  CA  501 GLUA 108 ...b    |   Z   3.2.1.1;
DDOOD    Ca  15   1   3   2  -1  -1  -1 ..... 6 1ava  1 0.2 1.9  CA  502 ASPA 127 .....   |    Z  3.2.1.1;
    1axn
OOOD     Ca   2   2  40  -1  -1  -1  -1 ....  4 1axn  0 0.4 1.8  CA  351 ILE   32 ...b    |
OOOE     Ca   3   2  40  -1  -1  -1  -1 ....  6 1axn  0 0.2 1.8  CA  353 GLY  187 ...b    |   ZZ
    1b0p
DTO      Mg  28   2  -1  -1  -1  -1  -1 ...   6 1b0p  0 0.3 2.3  MG 1237 ASPA 963 ...     |  UUZ  1.2.7.1;
DNOOO    Ca   2  71   3   2  -1  -1  -1 ..... 5 1b0p  2 0.4 2.3  CA 1238 ASPA 983 .....   |       1.2.7.1;
    1bag
```

```
NODO     Ca  36   9  34  -1  -1  -1  -1 ...   5 1bag  0 0.2 2.5  CA    6 ASN  101 ..b.    |   Z   3.2.1.1;
OD       Ca   2  -1  -1  -1  -1  -1  -1 ..    5 1bag  0 0.1 2.5  CA    7 GLY  169 .b      | ZZZ   3.2.1.1;
EO       Ca  37  -1  -1  -1  -1  -1  -1 ..    5 1bag  0 0.2 2.5  CA    8 GLU  276 b.      | ZZZ   3.2.1.1;
    1bfd
DNO      Ca  27   2  -1  -1  -1  -1  -1 ...   6 1bfd  0 0.2 1.6  CA  529 ASP  428 ...     |  UUZ  4.1.1.7;
OOO      Mg   1   2  -1  -1  -1  -1  -1 ...   3 1bfd  1 0.3 1.6  MG  531 ASN  117 ...     |       4.1.1.7;
    1c1y                                        ** dif chains  ***
ST       Mg  18  -1  -1  -1  -1  -1  -1 ..    6 1c1y  0 0.2 1.9  MG  171 SERA  17 ..      | UUZZ  2.7.1.-;
OE       Ca   2  -1  -1  -1  -1  -1  -1 ..    4 1c1y  0 0.2 1.9  CA  173 GLYB 123 .b      | ZZ    2.7.1.-;
    1c7k
HHD      Zn   4   6  -1  -1  -1  -1  -1 ee.   4 1c7k  0 0.1 1.0  ZN  133 HISA  83 ...     |  Z    3.4.24.-
DT       Ca   2  -1  -1  -1  -1  -1  -1 ..    6 1c7k  0 0.1 1.0  CA  134 ASPA  76 b.      | ZZZZ  3.4.24.-
    1cdl                                        ** ambiguities ***
DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  0 0.4 2.2  CA    1 ASPA  20 .b..b   |
DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  1 0.4 2.2  CA    2 ASPA  56 ....b   |
DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  0 0.3 2.2  CA    3 ASPA  93 .b..b   |
DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1cdl  1 0.6 2.2  CA    4 ASPA 129 .bb..   |
DDNNOE   Ca   2   2   0   2   5  -1  -1 ......6 1cdl  0 0.6 2.2  CA    2 ASPB  56 .....b  |
DDNODE   Ca   2   2   2   2   3  -1  -1 ......6 1cdl  0 0.6 2.2  CA    2 ASPC  56 .....b  |
    1clc
ONODD    Ca   3   2   2   3  -1  -1  -1 ..... 6 1clc  0 0.1 1.9  CA  591 GLU  236 .....   |    Z  3.2.1.4;
TODDO    Ca   2   3   1  39  -1  -1  -1 ..... 7 1clc  0 0.2 1.9  CA  592 THR  356 ..b..   |    ZZ 3.2.1.4;
ODO      Ca   3   2  -1  -1  -1  -1  -1 ...   6 1clc  0 0.1 1.9  CA  593 SER  520 ...     |  ZZZ  3.2.1.4;
CCHH     Zn  18   1  23  -1  -1  -1  -1 ..de  4 1clc  0 0.3 1.9  ZN  594 CYS  155 ....    |       3.2.1.4;
    1cru
EO       Ca  10  -1  -1  -1  -1  -1  -1 ..    6 1cru  0 0.1 1.5  CA  901 GLUB 253 b.      | ZZZZ  1.1.99.17;
OODE     Ca   2   2  36  -1  -1  -1  -1 ....  6 1cru  0 0.1 1.5  CA  902 ALAB 269 ...b    |   ZZ  1.1.99.17;
OO       Ca   1  -1  -1  -1  -1  -1  -1 ..    7 1cru  0 0.1 1.5  CA  908 GLYB 247 ..      | UBUZZ 1.1.99.17;
    1cvr
EHD      Ca   5 126  -1  -1  -1  -1  -1 .d.   6 1cvr  1 0.3 2.0  CA  648 GLUA 161 ...     |  ZZZ  3.4.22.37;
DOE      Ca 171   9  -1  -1  -1  -1  -1 ...   6 1cvr  0 0.2 2.0  CA  501 ASPA  78 b.b     |  ZZZ  3.4.22.37;
ODOE     Ca   3   2   2  -1  -1  -1  -1 ....  6 1cvr  0 0.2 2.0  CA  477 VALA 100 ...b    |   ZZ  3.4.22.37;
EEH      Ca   3   6  -1  -1  -1  -1  -1 ..e   4 1cvr  1 0.5 2.0  CA  686 GLUA 293 b..     |  Z    3.4.22.37;
EH       Zn  59  -1  -1  -1  -1  -1  -1 .e    3 1cvr  1 0.5 2.0  ZN  731 GLUA 152 ..      | Z     3.4.22.37;
    1d0b
OD       Ca   2  -1  -1  -1  -1  -1  -1 ..    7 1d0b  0 0.2 1.9  CA  201 PROA  49 ..      | ZZZZZ
ED       Ca   4  -1  -1  -1  -1  -1  -1 ..    5 1d0b  0 0.3 1.9  CA  202 GLUA  55 .b      | ZZZ
    1djx
NEDE     Ca  29   2  47  -1  -1  -1  -1 ....  6 1djx  0 0.3 2.3  CA    2 ASNA 312 ..b.    |   UZ  3.1.4.11;
ODN      Ca   2  24  -1  -1  -1  -1  -1 ...   5 1djx  1 0.4 2.3  CA    3 ILEA 651 .b.     |  ZZ   3.1.4.11;
    1dx5                                        ** dif chains ***
DOENOO   Ca   1   2  13   1   3  -1  -1 ......7 1dx5  0 0.3 2.3  CA 1001 ASPI 423 b.....  |     Z 3.4.21.5;
OO       Na   3  -1  -1  -1  -1  -1  -1 ..    5 1dx5  0 0.2 2.3  NA 2001 ARGN 221 ..      | ZZZ   3.4.21.5;
    1dyk
DOOD     Ca  17  49   2  -1  -1  -1  -1 ....  5 1dyk  0 0.2 2.0  CA 4001 ASPA2808 ....    |   Z
```

```
DOOD     Ca  17  54   2  -1  -1  -1  -1 ....  4 1dyk  0 0.2 2.0  CA 4002 ASPA2982 ....    |
    1e29
HH       Fe  51  -1  -1  -1  -1  -1  -1 ee    6 1e29  0 0.1 1.2  FE  136 HISA  41 ..      | BBBB
NN       Ca   1  -1  -1  -1  -1  -1  -1 ..    8 1e29  0 0.1 1.2  CA  225 ASNA  49 ..      | ZZZZZZ
OD       Ca   0  -1  -1  -1  -1  -1  -1 ..    6 1e29  0 0.3 1.2  CA  226 ASPA  35 ..      | ZZZZ
    1e43
NODDO    Ca  92   0   6  35  -1  -1  -1 ....  6 1e43  1 0.2 1.7  CA  501 ASNA 102 .....   |    Z  3.2.1.1;
DODDD    Ca  22   2  19   2  -1  -1  -1 ..... 6 1e43  0 0.1 1.7  CA  502 ASPA 159 b....   |    Z  3.2.1.1;
NE       Ca   3  -1  -1  -1  -1  -1  -1 ..    7 1e43  0 0.2 1.7  CA  503 ASNA 444 .b      | ZZZZZ 3.2.1.1;
OOODD    Ca   2 104   1  23  -1  -1  -1 .. .. 6 1e43  0 0.1 1.7  CA  504 GLYA 300 ....b   |    Z  3.2.1.1;
DDDDO    Na  24  11   6   1  -1  -1  -1 ..... 5 1e43  0 0.2 1.7  NA  505 ASPA 159 ..b..   |       3.2.1.1;
    1e8u                                        **  ambiguities ***
ODOOO    Ca   0   3   2  30  -1  -1  -1 ..... 5 1e8u  0 0.1 2.0  CA 1002 ASPA 261 .....   |       3.2.1.18
ODOSOO   Ca   0   3   0   2  30  -1  -1 ......6 1e8u  0 0.2 2.0  CA 1003 ASPB 261 ......  |       3.2.1.18
    1edm                                        ** dif chains  ***
OS       Ca   0  -1  -1  -1  -1  -1  -1 ..    7 1edm  0 0.2 1.5  CA    3 SERB  53 ..      | ZZZZZ
DOQDO    Ca   1   2  14   1  -1  -1  -1 ..... 5 1edm  0 0.2 1.5  CA    1 ASPC  47 ...b.   |
    1ezm
HHE      Zn   4  20  -1  -1  -1  -1  -1 ee.   4 1ezm  1 0.2 1.5  ZN  300 HIS  140 ...     |  Z    3.4.24.26) (
DEEDO    Ca  36   3   8   2  -1  -1  -1 ..... 6 1ezm  0 0.3 1.5  CA  400 ASP  136 ..b..   |    Z  3.4.24.26) (
    1fo4
OOOSSO   Ca   3   1   3  33   1  -1  -1 ......6 1fo4  0 0.4 2.1  CA 4009 ALAA 867 ......  |       1.1.1.204
CC       Fe  37  -1  -1  -1  -1  -1  -1 ..    5 1fo4  0 0.3 2.1  FE13001 CYSA 113 ..      | JXX   1.1.1.204
CC       Fe  32  -1  -1  -1  -1  -1  -1 ..    5 1fo4  0 0.3 2.1  FE23001 CYSA 116 ..      | JXX   1.1.1.204
CC       Fe  22  -1  -1  -1  -1  -1  -1 ..    5 1fo4  0 0.3 2.1  FE13002 CYSA  51 ..      | JXX   1.1.1.204
CC       Fe   5  -1  -1  -1  -1  -1  -1 ..    5 1fo4  0 0.3 2.1  FE23002 CYSA  43 ..      | JXX   1.1.1.204
    1fs7
HH       Fe 141  -1  -1  -1  -1  -1  -1 ee    6 1fs7  0 0.2 1.6  FE  509 HISA 172 ..      | BBBB
HH       Fe 113  -1  -1  -1  -1  -1  -1 ee    6 1fs7  0 0.1 1.6  FE  510 HISA 102 ..      | BBBB
HH       Fe 106  -1  -1  -1  -1  -1  -1 ee    6 1fs7  0 0.2 1.6  FE  511 HISA 299 ..      | BBBB
HH       Fe  42  -1  -1  -1  -1  -1  -1 ee    6 1fs7  0 0.1 1.6  FE  512 HISA 288 ..      | BBBB
EOOQ     Ca   1  56   2  -1  -1  -1  -1 ....  6 1fs7  0 0.1 1.6  CA  651 GLUA 217 b...    |   ZZ
    1fzc                                        ** ambiguities ***
DDO      Ca   2   2  -1  -1  -1  -1  -1 ...   5 1fzc  0 0.2 2.3  CA    2 ASPB 381 b..     |  ZZ
DDOO     Ca   2   2   2  -1  -1  -1  -1 ....  4 1fzc  0 0.2 2.3  CA    1 ASPC 318 b...    |
    1g0h
EDDD     Ca  16   3 117  -1  -1  -1  -1 ....  6 1g0h  0 0.3 2.3  CA  290 GLUA  65 ....    |   UU  3.1.3.25;
EDO      Ca  16   2  -1  -1  -1  -1  -1 ...   6 1g0h  0 0.2 2.3  CA  291 GLUA  65 b..     |  UZZ  3.1.3.25;
    1g4y
DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1g4y  0 0.2 1.6  CA 1001 ASPR  20 ....b   |    Z
DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 1g4y  0 0.4 1.6  CA 1002 ASPR  56 ....b   |
    1g5c                                        ** dif chains  ***
CHC      Zn  55   3  -1  -1  -1  -1  -1 .e.   4 1g5c  0 0.1 2.1  ZN 1001 CYSA  32 ...     |  Z    4.2.1.1;
EOO      Ca  21   3  -1  -1  -1  -1  -1 ...   6 1g5c  0 0.3 2.1  CA 1009 GLUF  97 b..     |  ZZZ  4.2.1.1;
OE       Ca   0  -1  -1  -1  -1  -1  -1 ..    5 1g5c  0 0.4 2.1  CA 1001 GLUC 147 ..      | ZZZ   4.2.1.1;
```

```
    1gcy
NDODO    Ca  35   3   8  35  -1  -1  -1 ..... 6 1gcy  0 0.2 1.6  CA  451 ASNA 116 .b...   |    Z  3.2.1.60;
DOOHDE   Ca   1  11   0   3   1  -1  -1 .. d..6 1gcy  0 0.2 1.6  CA  452 ASPA   1 b.....  |       3.2.1.60;
    1gen
OOOO     Ca  45  48  49  -1  -1  -1  -1 ....  7 1gen  0 0.3 2.2  CA  302 ASP  476 ....    |   ZZZ 3.4.24.24;
OOOO     Na  45  48  49  -1  -1  -1  -1 ....  6 1gen  0 0.3 2.2  NA  304 ILE  478 ....    |   JZ  3.4.24.24;
    1hdf
OOSD     Ca  27   2  41  -1  -1  -1  -1 ....  6 1hdf  0 0.2 2.3  CA 1101 LYSA  19 ....    |   ZZ
DOOS     Ca  17  28   2  -1  -1  -1  -1 ....  7 1hdf  0 0.3 2.3  CA 1102 ASPA  45 ....    |   ZZZ
    1hyo
DOOOT    Mg   1  19   3   1  -1  -1  -1 ..... 5 1hyo  0 0.4 1.3  MG 1004 ASPA 233 .....   |       3.7.1.2;
DEED     Ca  73   2  32  -1  -1  -1  -1 ....  6 1hyo  0 0.1 1.3  CA 1006 ASPA 126 ....    |   UU  3.7.1.2;
    1i76
OOOD     Ca  32   2   2  -1  -1  -1  -1 ....  6 1i76  0 0.1 1.2  CA  996 ASPA 137 ....    |   ZZ  3.4.24.34;
DOOODE   Ca   1   2   2  18   3  -1  -1 ......6 1i76  0 0.1 1.2  CA  997 ASPA 154 ......  |       3.4.24.34;
HDH      Zn   2  13  -1  -1  -1  -1  -1 e.e   3 1i76  0 0.1 1.2  ZN  998 HISA 147 ...     |       3.4.24.34;
HHH      Zn   4   6  -1  -1  -1  -1  -1 eee   5 1i76  0 0.3 1.2  ZN  999 HISA 197 ...     |  UU   3.4.24.34;
    1i8a
DNEDD    Ca   2   8   2   1  -1  -1  -1 ..... 6 1i8a  0 0.3 1.9  CA  190 ASPA  81 ...bb   |    Z  3.2.1.8;
ODODE    Ca   2   2   2 114  -1  -1  -1 ..... 6 1i8a  0 0.2 1.9  CA  191 VALA  10 ....b   |    Z  3.2.1.8;
DODDO    Ca   2  12  80   1  -1  -1  -1 ..... 6 1i8a  0 0.3 1.9  CA  192 ASPA  60 b.b..   |    Z  3.2.1.8;
    1iod                                        **  ambiguities ***
OSEEE    Ca   0   2   4  81  -1  -1  -1 ..... 6 1iod  0 0.3 2.3  CA  501 SERA  41 ...bb   |    Z  3.4.21.6
OSQEE    Ca   0   2   4  73  -1  -1  -1 ..... 7 1iod  1 0.3 2.3  CA  502 SERB 241 ....b   |    ZZ 3.4.21.6
    1kap
OOTDOD   Ca   2   2  28   2   3  -1  -1 ......6 1kap  1 0.1 1.6  CA  614 ARGP 253 ...b..  |       3.4.24.-;
ODOE     Ca   2  37   2  -1  -1  -1  -1 ....  6 1kap  0 0.1 1.6  CA  615 GLYP 288 ...b    |   ZZ  3.4.24.-;
OODOOD   Ca   2   2  13   2   3  -1  -1 ......6 1kap  0 0.3 1.6  CA  616 GLYP 334 .....b  |       3.4.24.-;
OODD     Ca   2   2  26  -1  -1  -1  -1 ....  6 1kap  1 0.2 1.6  CA  617 GLYP 370 ....    |   ZZ  3.4.24.-;
OONOOD   Ca   2   2  13   2   3  -1  -1 ......6 1kap  0 0.3 1.6  CA  618 ASNP 343 .....b  |       3.4.24.-;
OODOD    Ca   2   2  18   7  -1  -1  -1 ..... 6 1kap  0 0.1 1.6  CA  619 GLYP 361 .....   |    Z  3.4.24.-;
OODOOD   Ca   2   2  13   2   3  -1  -1 ......6 1kap  0 0.3 1.6  CA  620 GLYP 352 .....b  |       3.4.24.-;
DSDOD    Ca   2   2   2   2  -1  -1  -1 ... . 6 1kap  0 0.3 1.6  CA  621 ASPP 446 ....b   |    Z  3.4.24.-;
    1kit
OONDOT   Ca   3   0  33  24   0  -1  -1 ......6 1kit  0 0.2 2.3  CA  802 ALA  253 ...b..  |       3.2.1.18;
DDO      Ca  61   1  -1  -1  -1  -1  -1 ...   4 1kit  0 0.2 2.3  CA  803 ASP  621 bb.     |  Z    3.2.1.18;
    1nls
EDDH     Mn   2   9   5  -1  -1  -1  -1 ...e  6 1nls  0 0.1 0.9  MN  239 GLU    8 ....    |   ZZ
DOND     Ca   2   2   5  -1  -1  -1  -1 ....  6 1nls  0 0.1 0.9  CA  240 ASP   10 b...    |   ZZ
    1nps
OOSN     Ca  30   2  37  -1  -1  -1  -1 ....  5 1nps  0 0.6 1.8  CA   90 TYRA   7 ....    |   Z
NOS      Ca  17  26  -1  -1  -1  -1  -1 ...   6 1nps  0 0.5 1.8  CA  145 ASNA  36 ...     |  ZZZ
    1oac
 HHH     Cu  -1   2 163  -1  -1  -1  -1 .eed  4 1oac  0 0.1 2.0  CU  801 HISA 524 ....    |       1.4.3.6
DODDO    Ca   1   1 143   1  -1  -1  -1 ..... 6 1oac  0 0.1 2.0  CA  802 ASPA 533 .....   |    Z  1.4.3.6
```

```
EODE     Ca  94   3   2  -1  -1  -1  -1 ....  6 1oac  0 0.2 2.0  CA  803 GLUA 573 b...    |   ZZ  1.4.3.6
    1pa2
ODOODS   Ca   0   3   2   2   2  -1  -1 ......7 1pa2  0 0.1 1.5  CA  307 ASPA  43 ......  |     Z 1.11.1.7;
OTDOTOD  Ca   0  51   3   0   3   2  -1 ......7 1pa2  0 0.1 1.5  CA  308 THRA 170 ....... |       1.11.1.7;
    1pyt                                        ** dif chains ***
EOOQE    Ca   2   3   2   3  -1  -1  -1 ..... 5 1pyt  1 0.3 2.3  CA  650 GLUC 470 .....   |       3.4.17.1;
HEH      Zn   3 124  -1  -1  -1  -1  -1 d.d   4 1pyt  0 0.4 2.3  ZN  350 HISB  69 .b.     |  Z    3.4.17.1;
    1qho
DODEE    Ca   1   2  22   1  -1  -1  -1 ..... 6 1qho  1 0.2 1.7  CA  696 ASPA  76 ....b   |    Z  3.2.1.133;
NODO     Ca  53  14  34  -1  -1  -1  -1 ...   7 1qho  0 0.2 1.7  CA  697 ASNA 131 ..b.    |   ZZZ 3.2.1.133;
DONNOD   Ca   2   3   1  21   2  -1  -1 ......7 1qho  0 0.2 1.7  CA  698 ASPA  21 ......  |     Z 3.2.1.133;
    1qla                                        ** dif chains / ?  ***
HH       Fe  89  -1  -1  -1  -1  -1  -1 ee    6 1qla  0 0.1 2.2  FE    1 HISC  93 ..      | BBBB  1.3.99.1;
HH       Fe  99  -1  -1  -1  -1  -1  -1 ee    6 1qla  0 0.1 2.2  FE    2 HISC  44 ..      | BBBB  1.3.99.1;
CC       Fe  12  -1  -1  -1  -1  -1  -1 ..    5 1qla  0 0.3 2.2  FE1   3 CYSB  65 ..      | JXX   1.3.99.1;
CC       Fe   5  -1  -1  -1  -1  -1  -1 ..    5 1qla  0 0.3 2.2  FE2   3 CYSB  57 ..      | JXX   1.3.99.1;
OOOOO    Ca   1   1  20   2  -1  -1  -1 ..... 6 1qla  0 0.2 2.2  CA    9 SERA 371 .....   |    Z  1.3.99.1;
    1sac
DNEOD    Ca   1  77   1   1  -1  -1  -1 ..... 6 1sac  0 0.3 2.0  CA    1 ASPA  58 b.b..   |    O
EDQ      Ca   2  10  -1  -1  -1  -1  -1 ...   4 1sac  0 0.3 2.0  CA    2 GLUA 136 .b.     |  U
    1scj
QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 1scj  0 0.2 2.0  CA  381 GLNA   2 .b....  |       3.4.21.62;
OOOT     Ca   2   3   0  -1  -1  -1  -1 ....  5 1scj  0 0.4 2.0  CA  382 ALAA 169 ....    |   Z   3.4.21.62;
    1sra
DODOE    Ca   3   2   2   5  -1  -1  -1 ..... 6 1sra  0 0.2 2.0  CA  301 ASP  222 ....b   |    Z
DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1sra  0 0.2 2.0  CA  302 ASP  257 ....b   |    Z
OOE      Ca   2   3  -1  -1  -1  -1  -1 ...   5 1sra  1 0.4 2.0  CA  303 PRO  241 ...     |  ZZ
    1svy
DOO      Ca   5   2  -1  -1  -1  -1  -1 ...   6 1svy  0 0.3 1.8  CA    1 ASP  222 ...     |  ZZZ
ODE      Na   1  22  -1  -1  -1  -1  -1 ...   4 1svy  1 0.2 1.8  NA    2 GLY  187 ..b     |  Z
    1tf4
SODEO    Ca   1   3   1  46  -1  -1  -1 ..... 5 1tf4  0 0.2 1.9  CA 3001 SERA 210 ..bb.   |       3.2.1.4;
ODOND    Ca   2  65   3   1  -1  -1  -1 ..... 6 1tf4  0 0.2 1.9  CA 3002 THRA 504 .b...   |    Z  3.2.1.4;
    1tn3
DEOON    Ca   4  27   3   1  -1  -1  -1 ..... 6 1tn3  0 0.2 2.0  CA  182 ASP  116 bb...   |    Z
QEOD     Ca   7  15   0  -1  -1  -1  -1 ....  5 1tn3  2 0.2 2.0  CA  183 GLN  143 ....    |   Z
    1vrk
DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.1 1.9  CA  151 ASPA  20 ....b   |    Z  2.7.1.117;
DDNOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.1 1.9  CA  152 ASPA  56 ....b   |    Z  2.7.1.117;
DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 1vrk  0 0.2 1.9  CA  154 ASPA 129 ....b   |    Z  2.7.1.117;
    1wdc                                        ** dif chains  ***
ODDODO   Ca   0   3   1   2   2  -1  -1 ......7 1wdc  0 0.2 2.0  CA  501 ASPC  19 ......  |     Z
DDDOD    Mg   2   2   2   5  -1  -1  -1 ..... 6 1wdc  1 0.3 2.0  MG  502 ASPB  28 .....   |    Z
    2msb
DEDOD    Ca   4  23   5   1  -1  -1  -1 ..... 6 2msb  0 0.2 1.7  CA    1 ASPA 161 bb...   |    Z
```

```
ENENOD   Ca   2   6  12   1   0  -1  -1 ......6 2msb  0 0.2 1.7  CA    2 GLUA 185 ......  |
ED       Ca  29  -1  -1  -1  -1  -1  -1 ..    6 2msb  0 0.2 1.7  CA    3 GLUA 165 .b      | ZZZZ
    2por
ED       Ca  28  -1  -1  -1  -1  -1  -1 ..    6 2por  0 0.1 1.8  CA  302 GLU   80 b.      | ZZZZ
DDND     Ca   2   5   1  -1  -1  -1  -1 ....  6 2por  0 0.2 1.8  CA  303 ASP   93 bb..    |   ZZ
NDOO     Ca  20   2   2  -1  -1  -1  -1 ....  5 2por  0 0.2 1.8  CA  304 ASN  116 .b..    |   Z
    2pvb
DDSOEE   Ca   2   2   2   2   3  -1  -1 ......6 2pvb  0 0.1 0.9  CA  110 ASPA  51 .....b  |
DDDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 2pvb  0 0.1 0.9  CA  111 ASPA  90 ....b   |    Z
    2sas
DNDOD    Ca   2   2   2   5  -1  -1  -1 ..... 5 2sas  0 0.2 2.4  CA  186 ASP   19 ....b   |
DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 2sas  0 0.4 2.4  CA  187 ASP   70 ....b   |    Z
DSDOE    Ca   2   2   2   5  -1  -1  -1 ..... 5 2sas  0 0.2 2.4  CA  188 ASP  115 ....b   |
    2scp
DDDOD    Ca   2   2   2   5  -1  -1  -1 ..... 6 2scp  0 0.1 2.0  CA  190 ASPA  16 ....b   |    Z
DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 2scp  0 0.1 2.0  CA  191 ASPA 104 ....b   |    Z
DNDOE    Ca   2   2   2   5  -1  -1  -1 ..... 6 2scp  0 0.2 2.0  CA  192 ASPA 138 ....b   |    Z
    2sic
QDONOO   Ca  39  34   2   2   2  -1  -1 ......6 2sic  0 0.2 1.8  CA  501 GLNE   2 .b....  |       3.4.21.14) C
OOOD     Ca   5  21   2  -1  -1  -1  -1 ....  6 2sic  1 0.6 1.8  CA  502 GLYE 169 ....    |   ZZ  3.4.21.14) C
    1g8k
CC       Fe  18  -1  -1  -1  -1  -1  -1 ..    5 1g8k  0 0.3 1.6  FE15006 CYSB  60 ..      | JXX
HH       Fe  19  -1  -1  -1  -1  -1  -1 dd    5 1g8k  0 0.3 1.6  FE25006 HISB  62 ..      | JXX
    1a49                                        **  ambiguities ***
NSDOS    K    2  36   1 129  -1  -1  -1 ..... 6 1a49  0 0.3 2.1   K  532 ASNA  74 .....   |    U  2.7.1.40;
ED       Mg  24  -1  -1  -1  -1  -1  -1 ..    6 1a49  0 0.2 2.1  MG  534 GLUA 271 ..      | UUUZ  2.7.1.40;
NSODO    K    2  36   0   1  -1  -1  -1 ..... 7 1a49  0 0.4 2.1   K 1132 ASNB 674 .....   |    ZZ 2.7.1.40;
NNSDO    K    0   2  36   1  -1  -1  -1 ..... 7 1a49  1 0.4 2.1   K 4732 ASNG4274 .....   |    UZ 2.7.1.40;
NSDO     K    2  36   1  -1  -1  -1  -1 ....  6 1a49  0 0.1 2.1   K 5332 ASNH4874 ....    |   ZZ  2.7.1.40;
    1a73
CCHC     Zn   7   2   4  -1  -1  -1  -1 ..d.  4 1a73  0 0.2 1.8  ZN  201 CYSA 125 ....    |
CCCH     Zn  59   5   5  -1  -1  -1  -1 ...d  4 1a73  0 0.1 1.8  ZN  202 CYSA  41 ....    |
    1alo
CC       Fe  39  -1  -1  -1  -1  -1  -1 ..    5 1alo  0 0.2 2.0  FE1 907 CYS  100 ..      | JXX
CC       Fe  34  -1  -1  -1  -1  -1  -1 ..    5 1alo  0 0.2 2.0  FE2 907 CYS  103 ..      | JXX
CC       Fe   5  -1  -1  -1  -1  -1  -1 ..    5 1alo  0 0.3 2.0  FE1 908 CYS   40 ..      | JXX
CC       Fe  12  -1  -1  -1  -1  -1  -1 ..    5 1alo  0 0.3 2.0  FE2 908 CYS   48 ..      | JXX
OEO      Mg   2  42  -1  -1  -1  -1  -1 ...   4 1alo  0 0.3 2.0  MG  916 ALA  649 .b.     |  Z
EE       Mg   4  -1  -1  -1  -1  -1  -1 ..    5 1alo  0 0.4 2.0  MG  918 GLU  899 ..      | ZZZ
    1bpy
DD       Mg   2  -1  -1  -1  -1  -1  -1 ..    6 1bpy  0 0.3 2.2  MG  339 ASPA 190 ..      | UUUZ  2.7.7.7;
DDD      Mg   2  64  -1  -1  -1  -1  -1 ...   4 1bpy  0 0.3 2.2  MG  340 ASPA 190 b..     |  U    2.7.7.7;
OOO      Na   2   3  -1  -1  -1  -1  -1 ...   6 1bpy  0 0.5 2.2  NA  341 THRA 101 ...     |  UZZ  2.7.7.7;
OOO      Na   2   3  -1  -1  -1  -1  -1 ...   6 1bpy  0 0.4 2.2  NA  342 LYSA  60 ...     |  UZZ  2.7.7.7;
    1bup
```

```
DO       K    5  -1  -1  -1  -1  -1  -1 ..    7 1bup  1 0.3 1.7   K  490 ASPA  10 b.      | UUUZZ
ODOTD    K    0   5   0   2  -1  -1  -1 ..... 7 1bup  0 0.3 1.7   K  491 ASPA 199 .....   |    UZ
    1dp0                                        **  ambiguities ***
EHE      Mg   2  43  -1  -1  -1  -1  -1 .d.   6 1dp0  0 0.2 1.7  MG 3001 GLUA 416 ...     |  ZZZ  3.2.1.23;
OOOQD    Mg   3   3 142  30  -1  -1  -1 ..... 5 1dp0  0 0.4 1.7  MG 3002 ASPA  15 ....b   |       3.2.1.23;
DON      Na 400   3  -1  -1  -1  -1  -1 ...   5 1dp0  0 0.2 1.7  NA 3101 ASPA 201 ...     |  ZZ   3.2.1.23;
OOO      Na   3   3  -1  -1  -1  -1  -1 ...   5 1dp0  1 0.2 1.7  NA 3102 PHEA 556 ...     |  ZZ   3.2.1.23;
OOO      Na  35   3  -1  -1  -1  -1  -1 ...   5 1dp0  1 0.3 1.7  NA 3103 PROA 932 ...     |  ZZ   3.2.1.23;
OOO      Na   3  20  -1  -1  -1  -1  -1 ...   6 1dp0  1 0.3 1.7  NA 3104 SERA 647 ...     |  OZZ  3.2.1.23;
OOOO     Na   3   1   2  -1  -1  -1  -1 ....  6 1dp0  0 0.4 1.7  NA 3102 PHEC 556 ....    |   ZZ  3.2.1.23;
    1ew2
DHH      Zn   4 112  -1  -1  -1  -1  -1 .ee   5 1ew2  0 0.4 1.8  ZN 1001 ASPA 316 b..     |  UU   3.1.3.1
DSD      Zn  50 265  -1  -1  -1  -1  -1 ...   4 1ew2  1 0.4 1.8  ZN 1002 ASPA  42 b..     |  U    3.1.3.1
DSE      Mg 113 156  -1  -1  -1  -1  -1 ...   6 1ew2  0 0.3 1.8  MG 1003 ASPA  42 ...     |  ZZZ  3.1.3.1
EOED     Mg  53   1  15  -1  -1  -1  -1 ....  5 1ew2  0 0.3 1.8  MG 1004 GLUA 216 b..b    |   Z   3.1.3.1
    1ewk                                        **  ambiguities ***
OOOO     Mg   3   3   1  -1  -1  -1  -1 ....  6 1ewk  0 0.4 2.2  MG 1001 ILEA  89 ....    |   ZZ
OOO      Mg   6   1  -1  -1  -1  -1  -1 ...   5 1ewk  1 0.3 2.2  MG 1002 ILEB  89 ...     |  ZZ
    1eyz
EE       Mg  12  -1  -1  -1  -1  -1  -1 ..    5 1eyz  0 0.2 1.8  MG  402 GLUA 267 b.      | UUZ   2.1.2.-;
NOO      Na   1   2  -1  -1  -1  -1  -1 ...   6 1eyz  0 0.3 1.8  NA  961 ASNA 100 ...     |  ZZZ  2.1.2.-;
    1g29                                        ** dif chains / ? ambiguities ***
DYD      Mg   2  61  -1  -1  -1  -1  -1 ...   3 1g29  0 0.6 1.9  MG   16 ASP2 293 ...     |
ED       Mg   1  -1  -1  -1  -1  -1  -1 ..    3 1g29  1 0.3 1.9  MG   79 GLU2 292 ..      | J
EDK      Mg   1  66  -1  -1  -1  -1  -1 ...   4 1g29  0 0.5 1.9  MG   83 GLU1 292 ...     |  J
KE       Na  42  -1  -1  -1  -1  -1  -1 ..    3 1g29  0 0.2 1.9  NA  112 LYS1  11 ..      | Z
OE       Na   0  -1  -1  -1  -1  -1  -1 ..    3 1g29  0 0.3 1.9  NA  225 GLU1 181 ..      | Z
TD       Mg 121  -1  -1  -1  -1  -1  -1 ..    2 1g29  0 0.5 1.9  MG  264 THR2  43 .b      |
DOO      Na  74   1  -1  -1  -1  -1  -1 ...   3 1g29  0 0.6 1.9  NA  422 ASP1  82 ...     |
    1gsa
DE       Mg   8  -1  -1  -1  -1  -1  -1 ..    6 1gsa  0 0.3 2.0  MG  319 ASP  273 ..      | UUUZ  6.3.2.3;
EN       Mg   2  -1  -1  -1  -1  -1  -1 ..    6 1gsa  1 0.4 2.0  MG  320 GLU  281 b.      | UXUZ  6.3.2.3;
    1h2r
EOH      Mg 436  54  -1  -1  -1  -1  -1 ..e   6 1h2r  0 0.1 1.4  MG 1005 GLUL  62 ...     |  ZZZ  1.12.2.1;
CC       Fe 465  -1  -1  -1  -1  -1  -1 ..    4 1h2r  2 0.2 1.4  FE 1004 CYSL  84 ..      | JX    1.12.2.1;
    1i74
HDD      Mn   4  63  -1  -1  -1  -1  -1 e..   5 1i74  1 0.3 2.2  MN  401 HISA   8 .b.     |  UZ   3.6.1.1;
DDHD     Mn  61  22  52  -1  -1  -1  -1 ..e.  6 1i74  1 0.2 2.2  MN  402 ASPA  14 ....    |   UZ  3.6.1.1;
    1iow
EN       Mg   2  -1  -1  -1  -1  -1  -1 ..    5 1iow  0 0.3 1.9  MG  330 GLU  270 b.      | UUZ   6.3.2.4;
DE       Mg  13  -1  -1  -1  -1  -1  -1 ..    6 1iow  0 0.2 1.9  MG  331 ASP  257 ..      | UUUZ  6.3.2.4;
    1pox
DNO      Mg  27   2  -1  -1  -1  -1  -1 ...   6 1pox  0 0.2 2.1  MG  610 ASPA 447 ...     |  UUZ  1.2.3.3) MUT
OQ       Na   3  -1  -1  -1  -1  -1  -1 ..    3 1pox 20 0.5 2.1  NA  614 META 452 ..      | Z     1.2.3.3) MUT
    1ryp                                        ** dif chains ***
```

```
TOOO     Mg 111   3   3  -1  -1  -1  -1 ....  5 1ryp  0 0.3 1.9  MG    1 THRA  17 ....    |   Z   3.4.99.46;
OOO      Mg   3   3  -1  -1  -1  -1  -1 . .   4 1ryp  0 0.3 1.9  MG    4 THRM 183 ...     |  Z    3.4.99.46;
OOO      Mg   3   3  -1  -1  -1  -1  -1 ...   5 1ryp  0 0.3 1.9  MG    5 ALAJ 166 ...     |  ZZ   3.4.99.46;
OOO      Mg   3   3  -1  -1  -1  -1  -1 ...   3 1ryp  0 0.5 1.9  MG    6 ILEH 163 ...     |       3.4.99.46;
    1t7p
DOD      Mg   1 178  -1  -1  -1  -1  -1 ...   6 1t7p  0 0.2 2.2  MG 4001 ASPA 475 ...     |  UUU  2.7.7.7;
DD       Mg 179  -1  -1  -1  -1  -1  -1 ..    5 1t7p  1 0.3 2.2  MG 4002 ASPA 475 ..      | UZZ   2.7.7.7;
    2occ                                        ** dif chains  ***
HH       Fe 317  -1  -1  -1  -1  -1  -1 ee    6 2occ  0 0.2 2.3  FE  515 HISA  61 ..      | BBBB  1.9.3.1;
HHH      Cu  50   1  -1  -1  -1  -1  -1 dee   4 2occ  1 0.2 2.3  CU  517 HISA 240 ...     |  U    1.9.3.1;
OEOO     Na   0   5 396  -1  -1  -1  -1 ....  4 2occ  0 0.1 2.3  NA  519 GLUA  40 ....    |       1.9.3.1;
HCCM     Cu  35   4   7  -1  -1  -1  -1 d...  5 2occ  0 0.3 2.3  CU  228 HISB 161 ....    |   J   1.9.3.1;
COCH     Cu   2   2   4  -1  -1  -1  -1 ...d  5 2occ  0 0.2 2.3  CU  229 CYSB 196 ....    |   J   1.9.3.1;
CCCC     Zn   2  20   3  -1  -1  -1  -1 ....  4 2occ  0 0.1 2.3  ZN   99 CYSF  60 ....    |       1.9.3.1;
    1a9x                                        ** dif chains / ? ambiguities ***
EN       Mn   2  -1  -1  -1  -1  -1  -1 ..    5 1a9x  0 0.3 1.8  MN 7901 GLUG6299 b.      | UUZ
QE       Mn  14  -1  -1  -1  -1  -1  -1 ..    6 1a9x  0 0.2 1.8  MN 7902 GLNG6285 ..      | UUUZ
ENOOOS   K   21   2   1   3   5  -1  -1 ......6 1a9x  0 0.2 1.8   K 7903 GLUG6215 ......  |
OEEO     K    1 172   1  -1  -1  -1  -1 ....  7 1a9x  1 0.3 1.8   K 7904 ALAG6126 ....    |   ZZZ
QE       Mn  12  -1  -1  -1  -1  -1  -1 ..    4 1a9x  1 0.3 1.8  MN 7911 GLNG6829 ..      | UU
EN       K    2  -1  -1  -1  -1  -1  -1 ..    4 1a9x  0 0.3 1.8   K 7912 GLUG6841 b.      | UU
EOOOS    K   22   1   3   5  -1  -1  -1 ..... 5 1a9x  0 0.3 1.8   K 7913 GLUG6761 .....   |
OOT      K   28   2  -1  -1  -1  -1  -1 ...   5 1a9x  0 0.3 1.8   K 7940 ASPG6084 ...     |  ZZ
OO       K   96  -1  -1  -1  -1  -1  -1  .    4 1a9x  0 0.4 1.8   K 7941 HISH7516 ..      | ZZ
ETNNQ    K   27  39   0   2  -1  -1  -1 ..... 8 1a9x  1 0.5 1.8   K 7942 GLUG6217 .....   |    UZZ
OTO      K    0   1  -1  -1  -1  -1  -1 ...   6 1a9x  0 0.4 1.8   K 7943 THRG6143 ...     |  ZZZ
OEEON    K    1 172   1   1  -1  -1  -1 ..... 8 1a9x  0 0.3 1.8   K 3904 ALAC2126 .....   |    ZZZ
QQE      Mn   0  12  -1  -1  -1  -1  -1 ...   6 1a9x  0 0.4 1.8  MN 3911 GLNC2829 ...     |  UUZ
OO#T     K   28   2   0  -1  -1  -1  -1 ....  5 1a9x  0 0.5 1.8   K 3940 ASPC2084 ....    |   Z
ETNQ     K   27  39   2  -1  -1  -1  -1 ....  6 1a9x  0 0.4 1.8   K 5942 GLUE4217 ....    |   UZ
    1az9
DHEE     Mn  83  29  23  -1  -1  -1  -1 .e..  6 1az9  1 0.3 2.0  MN  441 ASP  271 ....    |   ZZ  3.4.11.9;
DDE      Mn  11 135  -1  -1  -1  -1  -1 ...   5 1az9  2 0.2 2.0  MN  442 ASP  260 b..     |  ZZ   3.4.11.9;
    1d3v
HDDD     Mn  23   4 104  -1  -1  -1  -1 d...  7 1d3v  0 0.3 1.7  MN  500 HISA 101 ....    |   JUU 3.5.3.1;
DHDD     Mn   2 106   2  -1  -1  -1  -1 .d..  6 1d3v  1 0.3 1.7  MN  501 ASPA 124 ...b    |   UU  3.5.3.1;
    1e9g
DDD      Mn   5  32  -1  -1  -1  -1  -1 ...   6 1e9g  0 0.2 1.1  MN 2005 ASPB 115 ...     |  UZZ  3.6.1.1;
DD       Mn   5  -1  -1  -1  -1  -1  -1 ..    6 1e9g  0 0.1 1.1  MN 2008 ASPB 147 ..      | UUZZ  3.6.1.1;
    1eqj
DHH      Mn   4  55  -1  -1  -1  -1  -1 .ee   5 1eqj  0 0.3 1.7  MN  601 ASPA 403 b..     |  UU   5.4.2.1;
DSDH     Mn  50 382   1  -1  -1  -1  -1 ...e  4 1eqj  1 0.3 1.7  MN  701 ASPA  12 b...    |       5.4.2.1;
    1f3i
DE       Mn 229  -1  -1  -1  -1  -1  -1 ..    5 1f3i  1 0.4 2.3  MN  498 ASPA  97 ..      | UZZ
EE       Mn 235  -1  -1  -1  -1  -1  -1 ..    5 1f3i  1 0.3 2.3  MN  499 GLUA 110 b.      | ZZZ
```

```
    1f52
EEE      Mn  81   8  -1  -1  -1  -1  -1 ...   5 1f52  1 0.2 2.5  MN  469 GLUA 131 ...     |  ZZ   6.3.1.2;
EHE      Mn 140  88  -1  -1  -1  -1  -1 .d.   7 1f52  0 0.3 2.5  MN  470 GLUA 129 ...     |  UUZZ 6.3.1.2;
    1f5a
DDD      Mn   2  52  -1  -1  -1  -1  -1 ...   6 1f5a  0 0.3 2.5  MN 1001 ASPA 113 ...     |  BUZ  2.7.7.19;
DD       Mn   2  -1  -1  -1  -1  -1  -1 ..    5 1f5a  3 0.4 2.5  MN 1002 ASPA 113 ..      | UUZ   2.7.7.19;
    1ii7
DHDH     Mn   2  39 159  -1  -1  -1  -1 .e.e  6 1ii7  0 0.2 2.2  MN  403 ASPA   8 ....    |   UZ
DNHH     Mn  35  89  33  -1  -1  -1  -1 ..ed  6 1ii7  0 0.2 2.2  MN  404 ASPA  49 ....    |   UZ
    1ksi
HHH      Cu   2 159  -1  -1  -1  -1  -1 eed   5 1ksi  0 0.3 2.2  CU  650 HISA 442 ...     |  ZZ   1.4.3.6;
DODDO    Mn   1   1 139   1  -1  -1  -1 ..... 6 1ksi  0 0.3 2.2  MN  653 ASPA 451 .....   |    Z  1.4.3.6;
    1dqi                                    **  ambiguities ***
EHHHCH   Fe   2  25   6  64   3  -1  -1 .eee.d6 1dqi  0 0.2 1.7  FE  501 GLUA  14 ......  |
HHHCH    Fe  25   6  64   3  -1  -1  -1 eee.d 6 1dqi  0 0.3 1.7  FE  502 HISB  16 .....   |    Z
    1guq
CCHH     Zn   3  60  49  -1  -1  -1  -1 ..dd  4 1guq  0 0.1 1.8  ZN  350 CYSA  52 ....    |       2.7.7.10;
EHHH     Fe  99  15   2  -1  -1  -1  -1 .dee  4 1guq  0 0.2 1.8  FE  351 GLUA 182 b...    |       2.7.7.10;
OO       K   13  -1  -1  -1  -1  -1  -1 ..    7 1guq  0 0.2 1.8   K  353 ASNA 153 ..      | ZZZZZ 2.7.7.10;
    1mty
EEEH     Fe  65  34   3  -1  -1  -1  -1 ...d  6 1mty  1 0.3 1.7  FE    3 GLUD 144 ....    |   ZZ  1.14.13.25;
EEH      Fe  30   3  -1  -1  -1  -1  -1 ..d   6 1mty  1 0.3 1.7  FE    4 GLUD 114 ...     |  ZZZ  1.14.13.25;
    1r2f
DEHE     Fe  31   3  91  -1  -1  -1  -1 ..d.  4 1r2f  1 0.4 2.1  FE  400 ASPA  67 ...b    |       1.17.4.1;
EEEH     Fe  60  34   3  -1  -1  -1  -1 ...d  5 1r2f  0 0.3 2.1  FE  401 GLUA  98 .b..    |   Z   1.17.4.1;
    1afr
EEEH     Fe  53  33   3  -1  -1  -1  -1 ...d  4 1afr  0 0.3 2.4  FE  364 GLUA 143 .b..    |       1.14.99.6;
EEHE     Fe  38   3  83  -1  -1  -1  -1 ..d.  4 1afr  0 0.3 2.4  FE  365 GLUA 105 b...    |       1.14.99.6;
    1e5d
HEHD     Fe   2  65  19  -1  -1  -1  -1 e.e.  6 1e5d  0 0.3 2.5  FE1 404 HISA  79 ....    |   OU
DDH      Fe  82  61  -1  -1  -1  -1  -1 ..e   6 1e5d  1 0.3 2.5  FE2 404 ASPA  83 ...     |  OUZ
    1ute
DDYH     Fe  38   3 168  -1  -1  -1  -1 ...e  6 1ute  0 0.2 1.5  FE1 501 ASPA  14 ....    |   UO  3.1.3.2;
DNHH     Fe  39  95  35  -1  -1  -1  -1 ..ed  6 1ute  0 0.3 1.5  FE2 501 ASPA  52 ....    |   UO  3.1.3.2;
    2hmq
EHHHD    Fe  15   4  24   5  -1  -1  -1 .eee. 6 2hmq  0 0.2 1.7  FE1 101 GLUA  58 .....   |    O
HHED     Fe  29   4  48  -1  -1  -1  -1 ee..  6 2hmq  0 0.4 1.7  FE2 101 HISA  25 ....    |   OZ
    1eg9
HHD      Fe   5 149  -1  -1  -1  -1  -1 ee.   4 1eg9  0 0.2 1.6  FE  752 HISA 208 ..b     |  U    1.14.12.12;
HH       Fe  21  -1  -1  -1  -1  -1  -1 dd    5 1eg9  0 0.3 1.6  FE1 751 HISA  83 ..      | JXX   1.14.12.12;
CC       Fe  20  -1  -1  -1  -1  -1  -1 ..    5 1eg9  0 0.3 1.6  FE2 751 CYSA  81 ..      | JXX   1.14.12.12;
    1aoz
HCHM     Cu  62   5   5  -1  -1  -1  -1 d.d.  4 1aoz  0 0.4 1.9  CU  701 HISA 445 ....    |       1.10.3.3)
HHH      Cu 344  56  -1  -1  -1  -1  -1 eee   4 1aoz  0 0.1 1.9  CU2 702 HISA 106 ...     |  U    1.10.3.3)
HHH      Cu  42 404  -1  -1  -1  -1  -1 dee   4 1aoz  0 0.2 1.9  CU3 702 HISA  62 ...     |  U    1.10.3.3)
```

```
HH       Cu 388  -1  -1  -1  -1  -1  -1 ee    3 1aoz  0 0.1 1.9  CU4 703 HISA  60 ..      | U     1.10.3.3)
    1bt3
HHH      Cu  21   9  -1  -1  -1  -1  -1 eee   5 1bt3  0 0.4 2.5  CU2 500 HISA  88 ...     |  JU   1.10.3.1;
HHH      Cu   4  30  -1  -1  -1  -1  -1 eee   5 1bt3  0 0.4 2.5  CU3 500 HISA 240 ...     |  JU   1.10.3.1;
    1a65
HCH      Cu  56   5  -1  -1  -1  -1  -1 d.d   3 1a65  0 0.2 2.2  CU    1 HISA 396 ...     |       1.10.3.2;
HHHH     Cu 288   2  50  -1  -1  -1  -1 eeee  5 1a65  0 0.2 2.2  CU    2 HISA 111 ....    |   O   1.10.3.2;
HHH      Cu  43 344  -1  -1  -1  -1  -1 dee   3 1a65  0 0.2 2.2  CU    3 HISA  66 ...     |       1.10.3.2;
    1gof
YHH      Cu 224  85  -1  -1  -1  -1  -1 .ee   4 1gof  1 0.2 1.7  CU  700 TYR  272 ...     |  U    1.1.3.9) (PH
ODOOTOE  Na   3   2   3   0 104   1  -1 ......7 1gof  0 0.3 1.7  NA  702 LYS   29 ....... |       1.1.3.9) (PH
    1lla
HHH      Cu   4  27  -1  -1  -1  -1  -1 eee   3 1lla 20 0.1 2.2  CU  629 HIS  173 ...     |
HHH      Cu   4  36  -1  -1  -1  -1  -1 eee   3 1lla 20 0.2 2.2  CU  630 HIS  324 ...     |
OOOD     Na   3  68   0  -1  -1  -1  -1 ....  6 1lla  0 0.3 2.2  NA  631 SER  507 ....    |   ZZ
    1phm
HHH      Cu   1  64  -1  -1  -1  -1  -1 ddd   3 1phm  0 0.2 1.9  CU  357 HIS  107 ...     |       1.14.17.3;
HHM      Cu   2  70  -1  -1  -1  -1  -1 ee.   4 1phm  0 0.3 1.9  CU  358 HIS  242 ...     |  Z    1.14.17.3;
    1yai                                        **  ambiguities ***
HHH      Cu   2  78  -1  -1  -1  -1  -1 dee   3 1yai  2 0.2 1.9  CU  152 HISA  45 ...     |       1.15.1.1;
HHHD     Zn   9   9   3  -1  -1  -1  -1 ddd.  4 1yai  1 0.2 1.9  ZN  153 HISA  70 ....    |       1.15.1.1;
HHHH     Cu   2  23  55  -1  -1  -1  -1 deee  5 1yai  0 0.4 1.9  CU  152 HISC  45 ....    |   Z   1.15.1.1;
    2mta                                        ** dif chains  ***
HCH      Cu  39   3  -1  -1  -1  -1  -1 d.d   3 2mta  1 0.3 2.4  CU    0 HISA  53 ...     |       1.4.99.3) CO
HM       Fe  40  -1  -1  -1  -1  -1  -1 e.    6 2mta  0 0.1 2.4  FE  200 HISC  61 ..      | BBBB  1.4.99.3) CO
    1cyx
COCH     Cu   2   2   4  -1  -1  -1  -1 ...d  5 1cyx  0 0.3 2.3  CU1 201 CYS  207 ....    |   J   1.10.3.-;
HCC      Cu  35   4  -1  -1  -1  -1  -1 d..   4 1cyx  1 0.2 2.3  CU2 201 HIS  172 ...     |  J    1.10.3.-;
    2cua                                        ** dif chains  ***
HE       Zn   2  -1  -1  -1  -1  -1  -1 e.    2 2cua  1 0.5 1.6  ZN  172 HISB 117 ..      |       1.9.3.1;
HCCM     Cu  35   4   7  -1  -1  -1  -1 d...  5 2cua  0 0.3 1.6  CU1 169 HISA 114 ....    |   J   1.9.3.1;
COCH     Cu   2   2   4  -1  -1  -1  -1 ...d  5 2cua  0 0.4 1.6  CU2 169 CYSA 149 ....    |   J   1.9.3.1;
    1ah7
DHHD     Zn  14  49   4  -1  -1  -1  -1 .de.  5 1ah7  0 0.2 1.5  ZN  246 ASP   55 ....    |   Z   3.1.4.3;
#OHD     Zn   0  13 108  -1  -1  -1  -1 ..e.  5 1ah7  1 0.2 1.5  ZN  248 TRP    1 ....    |   Z   3.1.4.3;
    1ak0
DHHD     Zn  15  56   4  -1  -1  -1  -1 .de.  4 1ak0  0 0.5 1.8  ZN  271 ASP   45 ....    |       3.1.30.1
HHD      Zn  23   4  -1  -1  -1  -1  -1 ee.   4 1ak0  0 0.4 1.8  ZN  272 HIS  126 ..b     |  U    3.1.30.1
#OHD     Zn   0   5 114  -1  -1  -1  -1 ..e.  4 1ak0  0 0.4 1.8  ZN  273 TRP    1 ....    |       3.1.30.1
HE       Zn 166  -1  -1  -1  -1  -1  -1 e.    3 1ak0  0 0.3 1.8  ZN  274 HIS   15 .b      | Z     3.1.30.1
    1amp
DEH      Zn  35 104  -1  -1  -1  -1  -1 ..e   4 1amp  0 0.3 1.8  ZN  501 ASP  117 .b.     |  Z    3.4.11.10)
HDD      Zn  20  62  -1  -1  -1  -1  -1 e..   4 1amp  0 0.2 1.8  ZN  502 HIS   97 ..b     |  Z    3.4.11.10)
    1bf6
HHED     Zn   2 111 118  -1  -1  -1  -1 ee..  4 1bf6  0 0.3 1.7  ZN    1 HISA  12 ....    |
```

```
EHH      Zn  33  28  -1  -1  -1  -1  -1 .de   3 1bf6  0 0.2 1.7  ZN    2 GLUA 125 ...     |
    1cg2
DEH      Zn  35 209  -1  -1  -1  -1  -1 ..e   4 1cg2  0 0.2 2.5  ZN  500 ASPA 141 .b.     |  Z    3.4.17.11;
HDE      Zn  29  59  -1  -1  -1  -1  -1 e..   4 1cg2  0 0.2 2.5  ZN  501 HISA 112 ..b     |  Z    3.4.17.11;
    1dsz                                        ** equivalent chains probably ***
CCCC     Zn   3  14   3  -1  -1  -1  -1 ....  4 1dsz  0 0.1 1.7  ZN 1121 CYSA1135 ....    |
CCCC     Zn   6  10   3  -1  -1  -1  -1 ....  4 1dsz  0 0.1 1.7  ZN 1122 CYSA1171 ....    |
CCCC     Zn   3  14   3  -1  -1  -1  -1 ....  4 1dsz  0 0.1 1.7  ZN 1221 CYSB1235 ....    |
CCCC     Zn   6  10   3  -1  -1  -1  -1 ....  4 1dsz  0 0.1 1.7  ZN 1222 CYSB1271 ....    |
    1e3j
CCCC     Zn   3   3   8  -1  -1  -1  -1 ....  4 1e3j  0 0.2 2.3  ZN  901 CYSA  96 ....    |
CHE      Zn  25   1  -1  -1  -1  -1  -1 .e.   4 1e3j  0 0.2 2.3  ZN  902 CYSA  41 ...     |  Z
    1ete                                        ** dif chains  ***
EH       Zn  22  -1  -1  -1  -1  -1  -1 .e    4 1ete  0 0.4 2.2  ZN  135 GLUA  58 ..      | ZZ
DE       Zn  64  -1  -1  -1  -1  -1  -1 ..    2 1ete  1 0.1 2.2  ZN 1136 ASPB  14 ..      |
DE       Zn   2  -1  -1  -1  -1  -1  -1 ..    3 1ete  1 0.3 2.2  ZN 1137 ASPB  40 ..      | Z
    1fio
##H      Zn   1   0  -1  -1  -1  -1  -1 . d   3 1fio  0 0.2 2.1  ZN  501 META  30 ...     |
HH       Zn   3  -1  -1  -1  -1  -1  -1 dd    2 1fio  0 0.2 2.1  ZN  502 HISA  75 ..      |
    1hwt
CCCC     Zn   3   7   7  -1  -1  -1  -1 ....  5 1hwt  0 0.3 2.5  ZN    1 CYSC  64 ....    |   J
CCCC     Zn  17   3   9  -1  -1  -1  -1 ....  5 1hwt  0 0.3 2.5  ZN    2 CYSC  64 ....    |   J
    1hzy
HHD      Zn   2 244  -1  -1  -1  -1  -1 ee.   5 1hzy  0 0.2 1.3  ZN  401 HISA  55 ...     |  UZ   3.1.8.1;
HH       Zn  29  -1  -1  -1  -1  -1  -1 de    5 1hzy  0 0.2 1.3  ZN  402 HISA 201 ..      | UZZ   3.1.8.1;
NO       Na 116  -1  -1  -1  -1  -1  -1 ..    6 1hzy  1 0.5 1.3  NA  405 ASNA  38 ..      | ZZZZ  3.1.8.1;
    1i1i
HHE      Zn   4  25  -1  -1  -1  -1  -1 ee.   4 1i1i  0 0.4 2.3  ZN  701 HISP 474 ..b     |  Z    3.4.24.16;
HE       Zn   4  -1  -1  -1  -1  -1  -1 e.    3 1i1i  0 0.2 2.3  ZN  702 HISP 160 ..      | Z     3.4.24.16;
    1ile
CCCC     Zn   3 205   3  -1  -1  -1  -1 ....  5 1ile  0 0.7 2.5  ZN 1101 CYS  181 ....    |   Z
CCCC     Zn   3  38   2  -1  -1  -1  -1 ....  5 1ile  0 0.8 2.5  ZN 1102 CYS  461 ....    |   Z
    1j79
HH D     Zn   2  -1 232  -1  -1  -1  -1 ee .  5 1j79  0 0.2 1.7  ZN  400 HISA  16 .. .    | U Z   3.5.2.3;
 HH      Zn  -1  38  -1  -1  -1  -1  -1 .de   4 1j79  1 0.2 1.7  ZN  401 HISA 139 ...     |  Z    3.5.2.3;
    1j9y
HDE      Zn  32  10  -1  -1  -1  -1  -1 e..   4 1j9y  0 0.2 1.9  ZN 1003 HISA  79 ..b     |  Z    3.2.1.78;
RHDE     Zn   3  72  37  -1  -1  -1  -1 .d..  4 1j9y  0 0.4 1.9  ZN 1004 ARGA 208 ..b.    |       3.2.1.78;
    1kev                                        **  ambiguities ***
CHD      Zn  22  91  -1  -1  -1  -1  -1 .e.   3 1kev  1 0.1 2.0  ZN  353 CYSA  37 ...     |       1.1.1.2;
CHED     Zn  22   1  90  -1  -1  -1  -1 .e..  4 1kev  0 0.3 2.0  ZN  353 CYSB  37 ....    |       1.1.1.2;
    1lam
DODE     Zn  77   0   2  -1  -1  -1  -1 ....  6 1lam  0 0.4 1.6  ZN  488 ASP  255 ....    |   JZ  3.4.11.1
KDDE     Zn   5  18  61  -1  -1  -1  -1 ....  6 1lam  0 0.4 1.6  ZN  489 LYS  250 ....    |   JZ  3.4.11.1
OO       Zn  98  -1  -1  -1  -1  -1  -1 ..    3 1lam 22 0.7 1.6  ZN  490 THR  173 ..      | Z     3.4.11.1
```

```
    1ptq
HCCC     Zn  30   3  16  -1  -1  -1  -1 d...  4 1ptq  0 0.3 2.0  ZN    1 HIS  231 ....    |       2.7.1.-;
CCHC     Zn   3  22   3  -1  -1  -1  -1 ..d.  4 1ptq  0 0.3 2.0  ZN    2 CYS  244 ....    |       2.7.1.-;
    1qh5
HHHD     Zn   2  54  24  -1  -1  -1  -1 ede.  6 1qh5  0 0.3 1.5  ZN  261 HISA  54 ....    |   ZZ  3.1.2.6;
DHDH     Zn   1  75  39  -1  -1  -1  -1 .e.e  6 1qh5  0 0.3 1.5  ZN  262 ASPA  58 ....    |   XZ  3.1.2.6;
    1qtw
HHE      Zn  40  36  -1  -1  -1  -1  -1 ee.   4 1qtw  0 0.1 1.0  ZN  301 HISA  69 ...     |  Z    3.1.21.2;
HDH      Zn  47   2  -1  -1  -1  -1  -1 e.e   5 1qtw  0 0.2 1.0  ZN  302 HISA 182 .b.     |  ZZ   3.1.21.2;
EDHE     Zn  34  37  45  -1  -1  -1  -1 ..d.  5 1qtw  0 0.2 1.0  ZN  303 GLUA 145 ....    |   Z   3.1.21.2;
    1rmd
CHCH     Zn   4  23   2  -1  -1  -1  -1 .d.d  4 1rmd  0 0.1 2.1  ZN  117 CYS    2 ....    |
CCCC     Zn   3  17   3  -1  -1  -1  -1 ....  4 1rmd  0 0.1 2.1  ZN  118 CYS   26 ....    |
CHCC     Zn   2  18   3  -1  -1  -1  -1 .d..  4 1rmd  0 0.1 2.1  ZN  119 CYS   41 ....    |
CCHH     Zn   5  12   4  -1  -1  -1  -1 ..ee  4 1rmd  0 0.1 2.1  ZN  120 CYS   91 ....    |
    1sml
DHH      Zn   1 136  -1  -1  -1  -1  -1 .ee   5 1sml  0 0.2 1.7  ZN  268 ASPA  88 ...     |  ZZ   3.5.2.6;
HHH      Zn   2  74  -1  -1  -1  -1  -1 ede   4 1sml  1 0.2 1.7  ZN  269 HISA  84 ...     |  Z    3.5.2.6;
    1taf
ED       Zn   4  -1  -1  -1  -1  -1  -1 ..    3 1taf  0 0.4 2.0  ZN 2001 GLUB  31 .b      | Z
QQ       Zn   0  -1  -1  -1  -1  -1  -1 ..    4 1taf  0 0.5 2.0  ZN 2007 GLNB  45 ..      | ZZ
    1ush
DHDQ     Zn   2  41 170  -1  -1  -1  -1 .e..  5 1ush  0 0.2 1.7  ZN  600 ASP   41 ....    |   U   3.1.3.5;
DNHH     Zn  32 101  35  -1  -1  -1  -1 ..ed  6 1ush  0 0.3 1.7  ZN  601 ASP   84 ....    |   UU  3.1.3.5;
    1vfy
CCCC     Zn   3  27   3  -1  -1  -1  -1 ....  4 1vfy  0 0.1 1.1  ZN  300 CYSA 192 ....    |
CCCH     Zn   3  21   3  -1  -1  -1  -1 ...d  4 1vfy  0 0.1 1.1  ZN  301 CYSA 176 ....    |
    1zme
CCCC     Zn   3   7   6  -1  -1  -1  -1 ....  5 1zme  0 0.4 2.5  ZN    1 CYSC  34 ....    |   J
CCCC     Zn  16   3   7  -1  -1  -1  -1 ....  5 1zme  0 0.4 2.5  ZN    2 CYSC  34 ....    |   J
    3cao
EE       Zn   3  -1  -1  -1  -1  -1  -1 ..    2 3cao  1 0.4 1.6  ZN  114 GLUA  26 b.      |
HH       Fe  16  -1  -1  -1  -1  -1  -1 ee    6 3cao  0 0.1 1.6  FE  104 HISA  24 ..      | BBBB
HH       Fe  22  -1  -1  -1  -1  -1  -1 ee    6 3cao  0 0.1 1.6  FE  105 HISA  41 ..      | BBBB
HH       Fe  59  -1  -1  -1  -1  -1  -1 ee    6 3cao  0 0.1 1.6  FE  106 HISA  27 ..      | BBBB
HH       Fe  21  -1  -1  -1  -1  -1  -1 ee    6 3cao  0 0.1 1.6  FE  107 HISA  79 ..      | BBBB
    4mt2
CCCC     Zn   4   5   5  -1  -1  -1  -1 ....  4 4mt2  0 0.2 2.0  ZN   67 CYS   15 ....    |
CCCC     Zn   6   2  11  -1  -1  -1  -1 ....  4 4mt2  0 0.1 2.0  ZN   68 CYS    7 ....    |
    1cm5                                        ** ambiguities ***
OOEO     Na   2  46   1  -1  -1  -1  -1 ....  4 1cm5  0 0.1 2.3  NA 1056 ALAA 652 ....    |       2.3.1.54;
OEO      Na  46   1  -1  -1  -1  -1  -1 ...   3 1cm5  1 0.2 2.3  NA 1057 LEUB 654 ...     |       2.3.1.54;
    1d7u
OOTOO    Na   3   0   1   3  -1  -1  -1 ..... 5 1d7u  0 0.3 2.0  NA  435 ALAA  95 .....   |       4.1.1.64;
OSOOD    K    2 223   2   2  -1  -1  -1 ..... 6 1d7u  0 0.2 2.0   K  436 LEUA  78 .....   |    Z  4.1.1.64;
```

```
    1e39
HH       Fe  57  -1  -1  -1  -1  -1  -1 ee    6 1e39  0 0.2 1.8  FE  801 HISA  18 ..      | BBBB
HH       Fe  32  -1  -1  -1  -1  -1  -1 ee    6 1e39  0 0.2 1.8  FE  802 HISA   8 ..      | BBBB
HH       Fe  14  -1  -1  -1  -1  -1  -1 ee    6 1e39  0 0.2 1.8  FE  803 HISA  58 ..      | BBBB
HH       Fe  25  -1  -1  -1  -1  -1  -1 ee    6 1e39  0 0.2 1.8  FE  804 HISA  61 ..      | BBBB
OOOOO    Na   1   1  26   2  -1  -1  -1 ..... 6 1e39  0 0.3 1.8  NA  810 THRA 506 .....   |    Z
    1hx6                                        ** dif chains  ***
ON       Na   0  -1  -1  -1  -1  -1  -1 ..    5 1hx6  0 0.2 1.6  NA  703 ASNA 262 ..      | ZZZ
DND      Na   3   3  -1  -1  -1  -1  -1 ...   6 1hx6  0 0.1 1.6  NA  705 ASPA 143 ...     |  ZZZ
EE       Na 150  -1  -1  -1  -1  -1  -1 ..    6 1hx6  0 0.2 1.6  NA  706 GLUB 111 ..      | ZZZZ
    1eex
QEEQO    K   29  51  75  66  -1  -1  -1 ..... 7 1eex  0 0.5 1.7   K  603 GLNA 141 .....   |    UU 4.2.1.28;
OSEE     K    3   1  15  -1  -1  -1  -1 ....  6 1eex  0 0.2 1.7   K  604 GLYA 261 ..b.    |   ZZ  4.2.1.28;
    1k4c
OO       K    1  -1  -1  -1  -1  -1  -1 ..    3 1k4c 40 0.3 2.0   K 3001 GLYC  77 ..      | J
OO       K    1  -1  -1  -1  -1  -1  -1 ..    4 1k4c 40 0.3 2.0   K 3002 VALC  76 ..      | JJ
OO       K    1  -1  -1  -1  -1  -1  -1 ..    4 1k4c 40 0.3 2.0   K 3003 THRC  75 ..      | JJ
OT       K    0  -1  -1  -1  -1  -1  -1 ..    3 1k4c 40 0.3 2.0   K 3004 THRC  75 ..      | J
```

ARCHITECTURE OF METAL COORDINATION GROUPS - TABLE 5W (continued)

Metal coordination groups in which one metal ion is coordinated by more than one protein chain within the asymmetric unit.

Equivalent coordination groups involving duplicate protein chains within the asymmetric unit have been eliminated. The column headings have the same meaning as in Table 3 a) and Table 3D. When the *seqdif* is given as -99 this indicates that the two donors are in different protein chains; for example in the first coordination group Zn HHH 4 6 starts at his A142 and the same Zn atom is part of Zn OY 0 where O and Y belong to another protein chain.

```
cngroup       seqdif 1 to 7             his  cn     cn2 rms res  metal   startAA  carbi othdonors e.c.no

HHHOY    Zn   4   6 -99   0  -1  -1  -1 eee.. 5 1atl  0 0.1 1.8  ZN  401 HISA 142 .....   |       3.4.24.42;
DDON     Ca   3 -99  56  -1  -1  -1  -1 ....  4 1c8n  0 0.5 2.3  CA  302 ASPB 160 ....    |
ONDD     Ca  56 -99   3  -1  -1  -1  -1 ....  5 1c8n  0 0.5 2.3  CA  303 THRA 219 ....    |   Z
ODOTODS  Ca   0 -99   0   2   2   2  -1 ......7 1d2v  0 0.2 1.8  CA  600 ASPA  96 ....... |       1.11.1.7;
DOQDON   Ca   1   2  14   1 -99  -1  -1 ......6 1edm  0 0.1 1.5  CA    2 ASPB  47 ...b..  |
DE       Ca -99  -1  -1  -1  -1  -1  -1 ..    2 1hei  2 0.5 2.1  CA  632 ASPA 437 .b      |
OOTOOQ   Ca   2   0 188   5 -99  -1  -1 ......6 1jg8  0 0.5 1.8  CA  901 THRB  12 ......  |       4.1.2.5;
QOTOO    Ca -99   0 188   5  -1  -1  -1 ..... 5 1jg8  1 0.4 1.8  CA  903 GLNA 236 .....   |       4.1.2.5;
QOOTOO   Ca -99   2   0 188   5  -1  -1 ......6 1jg8  0 0.5 1.8  CA  904 GLNB 236 ......  |       4.1.2.5;
HHHOS    Zn   4   6 -99   0  -1  -1  -1 eee.. 5 1kap  0 0.2 1.6  ZN  613 HISP 176 .....   |       3.4.24.-;
OSO      Ca   5 -99  -1  -1  -1  -1  -1 ...   6 1qd6  0 0.3 2.1  CA    1 ARGC 147 ...     |  ZZZ  3.1.1.32;
OOS      Ca -99   5  -1  -1  -1  -1  -1 ...   4 1qd6  0 0.2 2.1  CA    2 SERC 106 ...     |  Z    3.1.1.32;
HDDM     Zn  12  63 -99  -1  -1  -1  -1 e...  4 1qq9  1 0.1 1.5  ZN  901 HISA  85 ....    |       3.4.11.-
DEHOM    Zn  35 115 -99   0  -1  -1  -1 ..e.. 5 1qq9  0 0.4 1.5  ZN  902 ASPA  97 .b...   |       3.4.11.-
DDD      Ca -99 -99  -1  -1  -1  -1  -1 ...   3 2mpr  0 0.2 2.4  CA    1 ASPA  78 ...     |
QODD     Ca 287  29 -99  -1  -1  -1  -1 ....  6 1g8k  0 0.1 1.6  CA 5008 GLNA 467 ...b    |   UU
DQOD     Ca -99 287  29  -1  -1  -1  -1 ....  6 1g8k  0 0.1 1.6  CA 5108 ASPA 129 b...    |   UU
HHHEH    Fe  40 -99  15  32  -1  -1  -1 eee.e 5 1dxr  0 0.2 2.0  FE  500 HISL 190 ...b.   |
EE       Mg -99  -1  -1  -1  -1  -1  -1 ..    3 1aih  0 0.3 2.5  MG  103 GLUB 270 .b      | Z
DDD      Mg -99 -99  -1  -1  -1  -1  -1 ...   6 1e2a  0 0.1 2.3  MG   57 ASPA  81 ...     |  ZZZ  2.7.1.69;
DD       Mg -99  -1  -1  -1  -1  -1  -1 ..    6 1em9  0 0.2 2.0  MG  306 ASPA  71 ..      | ZZZZ
OO       Na -99  -1  -1  -1  -1  -1  -1 ..    5 1eyz  1 0.2 1.8  NA  960 ALAA  53 ..      | ZZZ   2.1.2.-;
DE       Mg -99  -1  -1  -1  -1  -1  -1 ..    2 1g29  0 0.5 1.9  MG  203 ASP1  80 ..      |
ESS      Mg -99   3  -1  -1  -1  -1  -1 ...   6 1ryp  0 0.4 1.9  MG    2 GLUE 105 ...     |  ZZZ  3.4.99.46;
OOOD     Mg   3   3 -99  -1  -1  -1  -1 ....  4 1ryp  0 0.2 1.9  MG    7 ILEI 163 ....    |       3.4.99.46;
OOOO     Mg   3   3 -99  -1  -1  -1  -1 ....  4 1ryp  0 0.3 1.9  MG    8 ALAL 165 ....    |       3.4.99.46;
HDE      Mg   1 -99  -1  -1  -1  -1  -1 e..   3 2occ  0 0.3 2.3  MG  518 HISA 368 ...     |       1.9.3.1;
DDOE     Mg   4 -99   1  -1  -1  -1  -1 ....  6 1qh8  0 0.2 1.6  MG 3002 ASPB 349 ....    |   UU  1.18.6.1;
OEDD     Mg   1 -99   4  -1  -1  -1  -1 ....  6 1qh8  0 0.1 1.6  MG 3006 LYSB 106 ....    |   UU  1.18.6.1;
```

```
DD       Mn -99  -1  -1  -1  -1  -1  -1 ..    5 1cnz  1 0.5 1.8  MN  801 ASPA 251 ..      | UZZ
DDD      Mn -99   4  -1  -1  -1  -1  -1 ...   3 1cnz  1 0.5 1.8  MN  802 ASPA 227 ...     |
HDE      Fe -99   4  -1  -1  -1  -1  -1 e..   3 1qgh  0 0.3 2.3  FE  157 HISE  31 ...     |
DEH      Fe   4 -99  -1  -1  -1  -1  -1 ..e   3 1qgh  0 0.3 2.3  FE  157 ASPE  58 ...     |
DHD      Zn -99   4  -1  -1  -1  -1  -1 .e.   3 1hfe  0 0.2 1.6  ZN  500 ASPS  86 ...     |       1.18.99.1;
HH       Cu -99  -1  -1  -1  -1  -1  -1 ee    2 1aoz  0 0.1 1.9  CU  812 HISA 286 ..      |       1.10.3.3)
DHEE     Zn   3 -99   4  -1  -1  -1  -1 .e..  4 1b0n  0 0.3 1.9  ZN 1001 ASPA  63 ..b.    |
DKD      Zn  37 -99  -1  -1  -1  -1  -1 ...   3 1b0n  1 0.2 1.9  ZN 1003 ASPA  55 ...     |
HE       Zn -99  -1  -1  -1  -1  -1  -1 d.    5 1b0n  1 0.4 1.9  ZN 1005 HISA  66 ..      | UUZ
H  D     Zn  -1  -1 -99  -1  -1  -1  -1 d  .  4 1cg2  0 0.2 2.5  ZN  502 HISA 229 .  .    |ZZ     3.4.17.11;
DH       Zn -99  -1  -1  -1  -1  -1  -1 .d    4 1cg2  0 0.3 2.5  ZN  502 ASPA 387 ..      | ZZ    3.4.17.11;
DCHC     Zn   2  63 -99  -1  -1  -1  -1 ..e.  4 1d8d  0 0.3 2.0  ZN 1001 ASPB 297 b...    |       2.5.1.21;
HDCH     Zn -99   2  50  -1  -1  -1  -1 e..e  4 1dce  0 0.2 2.0  ZN  900 HISA   2 .b..    |       2.5.1.-;
HEHO     Zn   3 125 -99   0  -1  -1  -1 d.d.. 5 1dtd  2 0.6 1.6  ZN  301 HISA 424 .....   |       3.4.15.1;
EEHE     Zn  26   3 -99  -1  -1  -1  -1 ..d.  4 1ec5  0 0.1 2.5  ZN   50 GLUB  10 b...    |
EEEH     Zn -99  26   3  -1  -1  -1  -1 ...d  4 1ec5  0 0.2 2.5  ZN   50 GLUB  36 .b..    |
CCCC     Zn   5 -99   5  -1  -1  -1  -1 ....  4 1ed5  0 0.1 1.8  ZN  900 CYSA  96 ....    |       1.14.13.39;
HEQ      Zn  38 -99  -1  -1  -1  -1  -1 e..   4 1f83  1 0.5 2.0  ZN  500 HISA 229 .b.     |  Z    3.4.24.69;
HHHH     Zn  11 -99  11  -1  -1  -1  -1 eded  4 1hwt  0 0.2 2.5  ZN    9 HISD  80 ....    |
DD       Zn -99  -1  -1  -1  -1  -1  -1 ..    4 1taf  0 0.4 2.0  ZN 2003 ASPA  73 .b      | ZZ
CC       Zn -99  -1  -1  -1  -1  -1  -1 ..    2 3eip  1 0.4 1.8  ZN  285 CYSA  47 ..      |
HEHO     Zn   3 124 -99  -1  -1  -1  -1 d.d.  4 4cpa  0 0.3 2.5  ZN    1 HIS   69 .b..    |       3.4.17.1) CO
TODO     Na   1 -99   1  -1  -1  -1  -1 . ..  5 1f80  1 0.6 2.3  NA   12 THRA 104 ....    |   Z   2.7.8.7;
```