# W O R L D W I D E
# wwPDB
## P R O T E I N   D A T A   B A N K

# wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:43 PM GMT

| | | |
|---|---|---|
| PDB ID | : | 1QBK |
| Title | : | STRUCTURE OF THE KARYOPHERIN BETA2-RAN GPPNHP NUCLEAR TRANSPORT COMPLEX |
| Authors | : | Chook, Y.M.; Blobel, G. |
| Deposited on | : | 1999-04-23 |
| Resolution | : | 3.00 Å(reported) |

The following versions of software and data (see references (i)) were used in the production of this report:

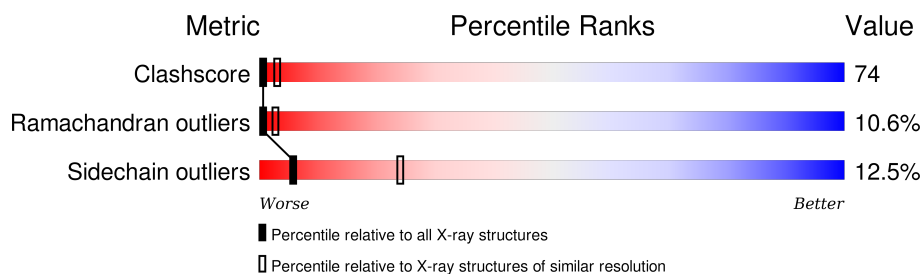| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.7 (RC4), CSD as536be (2015) |
| Xtriage (Phenix) | : | NOT EXECUTED |
| EDS | : | NOT EXECUTED |
| Percentile statistics | : | 20151230.v01 (using entries in the PDB archive December 30th 2015) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | trunk26865 |

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
*X-RAY DIFFRACTION*

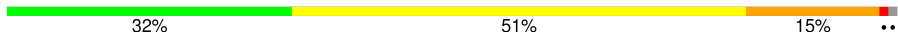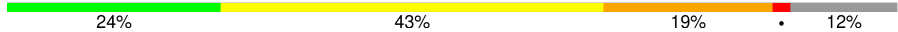The reported resolution of this entry is 3.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



| Metric | Percentile Ranks | Value |
| --- | --- | --- |
| Clashscore | | 74 |
| Ramachandran outliers | | 10.6% |
| Sidechain outliers | | 12.5% |

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
| --- | --- | --- |
| Clashscore | 102246 | 1912 (3.00-3.00) |
| Ramachandran outliers | 100387 | 1853 (3.00-3.00) |
| Sidechain outliers | 100360 | 1856 (3.00-3.00) |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Note EDS was not executed.

| Mol | Chain | Length | Quality of chain |
| --- | --- | --- | --- |
| 1 | B | 890 | 32% 51% 15% •• |
| 2 | C | 216 | 24% 43% 19% • 12% |

# 2   Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8572 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called KARYOPHERIN BETA2.

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | B | 880 | Total<br>7011 | C<br>4471 | N<br>1170 | O<br>1320 | S<br>26 | Se<br>24 | 0 | 0 | 0 |

There are 25 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| B | 217 | THR | ILE | CONFLICT | UNP Q92973 |
| B | 175 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 210 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 237 | ARG | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 247 | MET | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 254 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 261 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 308 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 429 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 432 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 482 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 564 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 566 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 575 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 624 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 641 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 672 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 680 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 710 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 739 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 743 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 749 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 789 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 820 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |
| B | 852 | MSE | MET | MODIFIED RESIDUE | UNP Q92973 |

- Molecule 2 is a protein called RAN.

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | C | 190 | Total<br>1528 | C<br>993 | N<br>265 | O<br>264 | S<br>3 | Se<br>3 | 0 | 0 | 0 |

There are 4 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| C | 129 | ARG | SER | CONFLICT | UNP P62826 |
| C | 89 | MSE | MET | MODIFIED RESIDUE | UNP P62826 |
| C | 179 | MSE | MET | MODIFIED RESIDUE | UNP P62826 |
| C | 189 | MSE | MET | MODIFIED RESIDUE | UNP P62826 |

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 3 | C | 1 | Total<br>1 | Mg<br>1 | 0 | 0 |

- Molecule 4 is PHOSPHOAMINOPHOSPHONIC ACID-GUANYLATE ESTER (three-letter code: GNP) (formula: $C_{10}H_{17}N_6O_{13}P_3$).
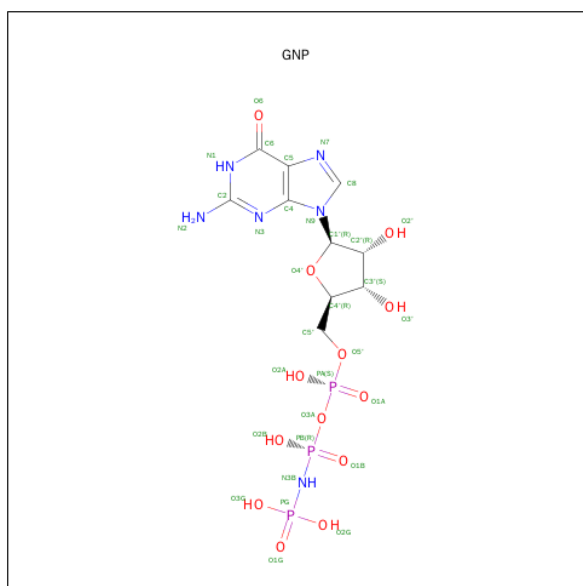


| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|---|
| 4 | C | 1 | Total<br>32 | C<br>10 | N<br>6 | O<br>13 | P<br>3 | 0 | 0 |

# 3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS was not executed.

- Molecule 1: KARYOPHERIN BETA2

Chain B:

Q744 P745 Y746 I747 P748 M749 V750 L751 H752 Q753 I757 I758 P764 K765 E769 C782 E785 M789 L790 Q791 Q792 F793 I794 R795 P796 W797 L801 R802 N803 I804 R805 D806 K810 D811 R815 G816 I817 M820 I821 N824 P825 S826 G827 V828 I829 Q830 F834

W842 I843 N844 D848 L849 R850 D851 I856 K861 N862 D866 R870 L878 P879 L880 K881 Y888 G889 V890

• Molecule 2: RAN

Chain C:

24%　　　43%　　　19%　　•　12%

MET ALA ALA GLY GLU PRO Q8 V9 Q10 F11 K12 L13 V14 L15 V16 G17 D18 G19 G20 T21 G22 K23 K28 R29 H30 L31 T32 G33 E34 F35 E36 K37 K38 Y39 V40 A41 T42 L43 G44 V45 E46 V47 H48 P49 L50 V51 T54 N55 R56 G57 P58 I59 K60 F61 N62 V63 W64

D65 T66 A67 G68 Q69 E70 K71 F72 G73 G74 L75 R76 D77 G78 Y79 Y80 I81 Q82 A83 Q84 C85 A86 I87 I88 M89 F90 D91 V92 T93 S94 R95 V96 T97 Y98 K99 N100 V101 P102 N103 W104 H105 R106 D107 L108 V109 R110 V111 C112 E113 N114 I115 P116 I117 V118 L119 N122 K123 V124

K127 D128 R129 K130 V131 K132 A133 K134 S135 I136 V137 F138 H139 R140 K141 K142 N143 L144 Q145 D148 I149 S150 A151 K152 S153 N154 Y155 N156 F157 L162 W163 L164 A165 R166 K167 I168 I169 L174 E175 F176 V177 A178 M179 P184 P185 E186 V187 V188 M189 D190 P191 A192 L193 A194 A195 Q196

Y197 GLU HIS ASP LEU GLU VAL ALA GLN THR THR ALA LEU PRO ASP GLU ASP ASP ASP LEU

# 4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

| Property | Value | Source |
|---|---|---|
| Space group | P 32 2 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 133.11Å   133.11Å   138.41Å<br>90.00°    90.00°    120.00° | Depositor |
| Resolution (Å) | 500.00  –  3.00 | Depositor |
| % Data completeness<br>(in resolution range) | 99.3 (500.00-3.00) | Depositor |
| $R_{merge}$ | 0.10 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| Refinement program | CNS | Depositor |
| R, $R_{free}$ | 0.267  ,  0.315 | Depositor |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 8572 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 77.0 | wwPDB-VP |

# 5 Model quality (i)

## 5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GNP, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Chain | Bond lengths | | Bond angles | |
|-----|-------|------|------------|------|---------|
| | | RMSZ | #$|Z|$>5 | RMSZ | #$|Z|$>5 |
| 1 | B | 0.56 | 2/7134 (0.0%) | 0.71 | 0/9651 |
| 2 | C | 0.49 | 0/1564 | 0.69 | 0/2113 |
| All | All | 0.55 | 2/8698 (0.0%) | 0.71 | 0/11764 |

All (2) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|-----|-------------|----------|
| 1 | B | 605 | CYS | CB-SG | -5.15 | 1.73 | 1.81 |
| 1 | B | 641 | MSE | CG-SE | -5.03 | 1.78 | 1.95 |

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | B | 7011 | 0 | 6991 | 1059 | 0 |
| 2 | C | 1528 | 0 | 1553 | 228 | 0 |
| 3 | C | 1 | 0 | 0 | 0 | 0 |
| 4 | C | 32 | 0 | 13 | 4 | 0 |
| All | All | 8572 | 0 | 8557 | 1265 | 0 |

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including

wwPDB X-ray Structure Validation Summary Report

hydrogen atoms). The all-atom clashscore for this structure is 74.

The worst 5 of 1265 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 2:C:83:ALA:HB3 | 2:C:112:CYS:SG | 1.55 | 1.46 |
| 1:B:143:ASN:HA | 1:B:148:ALA:O | 1.42 | 1.17 |
| 1:B:196:ALA:HB2 | 1:B:233:LYS:HB3 | 1.31 | 1.12 |
| 2:C:12:LYS:HG2 | 2:C:83:ALA:HA | 1.15 | 1.11 |
| 1:B:301:ILE:H | 1:B:302:PRO:HD2 | 1.17 | 1.09 |

There are no symmetry-related clashes.

## 5.3 Torsion angles (i)

### 5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles | |
|---|---|---|---|---|---|---|---|
| 1 | B | 874/890 (98%) | 604 (69%) | 185 (21%) | 85 (10%) | 1 | 3 |
| 2 | C | 188/216 (87%) | 130 (69%) | 30 (16%) | 28 (15%) | 0 | 1 |
| All | All | 1062/1106 (96%) | 734 (69%) | 215 (20%) | 113 (11%) | 0 | 3 |

5 of 113 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | B | 27 | THR |
| 1 | B | 50 | LEU |
| 1 | B | 81 | GLN |
| 1 | B | 83 | PHE |
| 1 | B | 103 | SER |

#### 5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles | |
|-----|-------|----------|-----------|----------|-------------|---|
| 1 | B | 790/778 (102%) | 697 (88%) | 93 (12%) | 6 | 26 |
| 2 | C | 164/182 (90%) | 138 (84%) | 26 (16%) | 3 | 15 |
| All | All | 954/960 (99%) | 835 (88%) | 119 (12%) | 6 | 24 |

5 of 119 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | B | 445 | GLN |
| 1 | B | 578 | ASP |
| 2 | C | 117 | ILE |
| 1 | B | 457 | ILE |
| 1 | B | 502 | SER |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | B | 537 | ASN |
| 1 | B | 632 | GLN |
| 2 | C | 103 | ASN |
| 1 | B | 554 | HIS |
| 1 | B | 610 | GLN |

#### 5.3.3 RNA (i)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|-----|------|-------|-----|------|--------|------|----------|--------|------|----------|
| | | | | | Counts | RMSZ | #$|Z| > 2$ | Counts | RMSZ | #$|Z| > 2$ |
| 4 | GNP | C | 218 | 3 | 28,34,34 | 1.87 | 5 (17%) | 33,54,54 | 2.31 | 6 (18%) |

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|-----|------|-------|-----|------|---------|----------|-------|
| 4 | GNP | C | 218 | 3 | - | 0/12/38/38 | 0/3/3/3 |

All (5) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|--------|------|-------------|----------|
| 4 | C | 218 | GNP | PG-O2G | -4.20 | 1.45 | 1.56 |
| 4 | C | 218 | GNP | C8-N7 | -3.79 | 1.27 | 1.34 |
| 4 | C | 218 | GNP | PB-O2B | -3.37 | 1.47 | 1.56 |
| 4 | C | 218 | GNP | O4'-C1' | 2.72 | 1.44 | 1.41 |
| 4 | C | 218 | GNP | C6-N1 | 4.90 | 1.42 | 1.33 |

The worst 5 of 6 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|--------|------|-------------|----------|
| 4 | C | 218 | GNP | C5-C6-N1 | -8.40 | 112.10 | 123.59 |
| 4 | C | 218 | GNP | N3-C2-N1 | -3.89 | 121.52 | 127.44 |
| 4 | C | 218 | GNP | O3G-PG-O2G | -2.35 | 100.60 | 107.58 |
| 4 | C | 218 | GNP | O3G-PG-O1G | -2.14 | 107.80 | 113.49 |

*Continued on next page...*

wwPDB X-ray Structure Validation Summary Report

*Continued from previous page...*

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|---|----------------|-------------|
| 4 | C | 218 | GNP | O3A-PA-O5' | -2.00 | 97.62 | 102.94 |

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 4 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 4 | C | 218 | GNP | 4 | 0 |

## 5.7 Other polymers (i)

There are no such residues in this entry.

## 5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

# 6 Fit of model and data (i)

## 6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section will therefore be empty.

## 6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section will therefore be empty.

## 6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

## 6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

## 6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.