



wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 03:02 AM GMT

PDB ID : 4J5W
Title : Crystal Structure of the apo-PXR/RXRalpha LBD Heterotetramer Complex
Authors : Wallace, B.D.; Betts, L.; Redinbo, M.R.
Deposited on : 2013-02-10
Resolution : 2.80 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

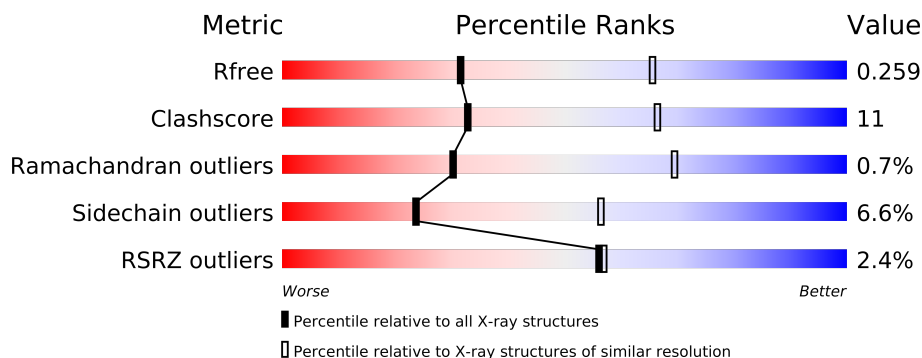
The following versions of software and data (see [references](#)) were used in the production of this report:

| | | |
|--------------------------------|---|--------------------------|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.15 2013 |
| Xtriage (Phenix) | : | dev-1323 |
| EDS | : | stable22639 |
| Percentile statistics | : | 21963 |
| Refmac | : | 5.8.0049 |
| CCP4 | : | 6.3.0 (Settle) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et. al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | stable22683 |

1 Overall quality at a glance

The reported resolution of this entry is 2.80 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 | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|-----------------------|-----------------------------|---|
| R_{free} | 66092 | 1799 (2.80-2.80) |
| Clashscore | 79885 | 2295 (2.80-2.80) |
| Ramachandran outliers | 78287 | 2252 (2.80-2.80) |
| Sidechain outliers | 78261 | 2254 (2.80-2.80) |
| RSRZ outliers | 66119 | 1802 (2.80-2.80) |

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | A | 336 | |
| 1 | B | 336 | |
| 2 | C | 264 | |
| 2 | D | 264 | |

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

| Mol | Type | Chain | Res | Geometry | Electron density |
|-----|------|-------|-----|----------|------------------|
| 3 | MG | D | 501 | - | X |

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8580 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Nuclear receptor subfamily 1 group I member 2, Nuclear receptor coactivator 1.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|---------|-------|
| 1 | A | 294 | Total | C | N | O | S | 0 | 1 | 0 |
| | | | 2406 | 1543 | 417 | 428 | 18 | | | |
| 1 | B | 293 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 2383 | 1527 | 413 | 425 | 18 | | | |

There are 16 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|----------------|------------|
| A | 127 | SER | - | EXPRESSION TAG | UNP O75469 |
| A | 128 | ASN | - | EXPRESSION TAG | UNP O75469 |
| A | 129 | ALA | - | EXPRESSION TAG | UNP O75469 |
| A | 435 | GLY | - | LINKER | UNP Q15788 |
| A | 436 | GLY | - | LINKER | UNP Q15788 |
| A | 437 | SER | - | LINKER | UNP Q15788 |
| A | 438 | GLY | - | LINKER | UNP Q15788 |
| A | 439 | GLY | - | LINKER | UNP Q15788 |
| B | 127 | SER | - | EXPRESSION TAG | UNP O75469 |
| B | 128 | ASN | - | EXPRESSION TAG | UNP O75469 |
| B | 129 | ALA | - | EXPRESSION TAG | UNP O75469 |
| B | 435 | GLY | - | LINKER | UNP Q15788 |
| B | 436 | GLY | - | LINKER | UNP Q15788 |
| B | 437 | SER | - | LINKER | UNP Q15788 |
| B | 438 | GLY | - | LINKER | UNP Q15788 |
| B | 439 | GLY | - | LINKER | UNP Q15788 |

- Molecule 2 is a protein called Retinoic acid receptor RXR-alpha, Nuclear receptor coactivator 1.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|---------|-------|
| 2 | C | 226 | Total | C | N | O | S | 0 | 1 | 0 |
| | | | 1804 | 1157 | 315 | 322 | 10 | | | |

Continued on next page...

Continued from previous page...

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|---------|-------|
| 2 | D | 227 | Total | C | N | O | S | 0 | 1 | 0 |
| | | | 1810 | 1161 | 316 | 323 | 10 | | | |

There are 10 discrepancies between the modelled and reference sequences:

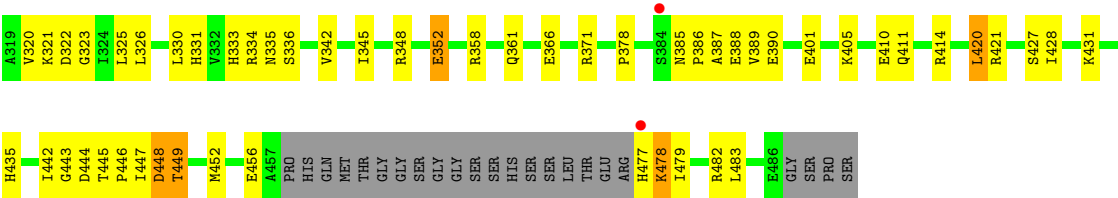
| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|------------|
| C | 463 | GLY | - | LINKER | UNP P19793 |
| C | 464 | GLY | - | LINKER | UNP P19793 |
| C | 465 | SER | - | LINKER | UNP P19793 |
| C | 466 | GLY | - | LINKER | UNP Q15788 |
| C | 467 | GLY | - | LINKER | UNP Q15788 |
| D | 463 | GLY | - | LINKER | UNP P19793 |
| D | 464 | GLY | - | LINKER | UNP P19793 |
| D | 465 | SER | - | LINKER | UNP P19793 |
| D | 466 | GLY | - | LINKER | UNP Q15788 |
| D | 467 | GLY | - | LINKER | UNP Q15788 |

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

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|----|---------|---------|
| 3 | D | 1 | Total | Mg | 0 | 0 |
| | | | 1 | 1 | | |
| 3 | C | 1 | Total | Mg | 0 | 0 |
| | | | 1 | 1 | | |

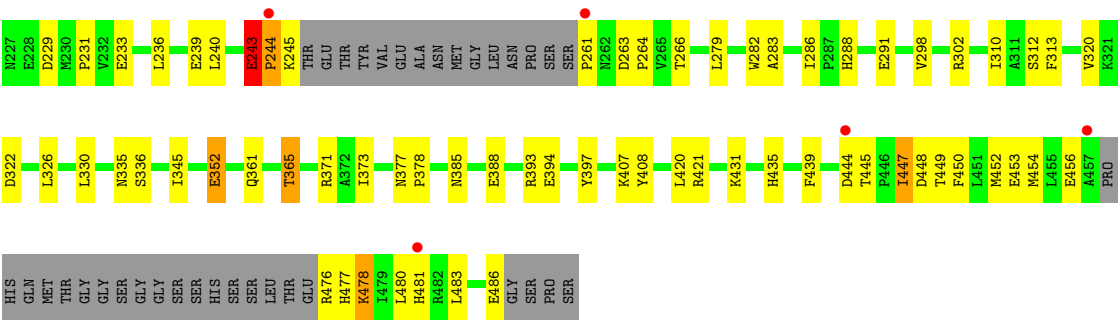
- Molecule 4 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|----|---------|---------|
| 4 | A | 55 | Total | O | 0 | 0 |
| | | | 55 | 55 | | |
| 4 | B | 44 | Total | O | 0 | 0 |
| | | | 44 | 44 | | |
| 4 | C | 26 | Total | O | 0 | 0 |
| | | | 26 | 26 | | |
| 4 | D | 50 | Total | O | 0 | 0 |
| | | | 50 | 50 | | |



● Molecule 2: Retinoic acid receptor RXR-alpha, Nuclear receptor coactivator 1

Chain D:



4 Data and refinement statistics

| Property | Value | Source |
|---|---|------------------|
| Space group | P 21 21 21 | Depositor |
| Cell constants a, b, c, α , β , γ | 70.26Å 109.55Å 169.88Å 90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 48.54 – 2.80 48.54 – 2.80 | Depositor EDS |
| % Data completeness (in resolution range) | 96.0 (48.54-2.80) 96.1 (48.54-2.80) | Depositor EDS |
| R_{merge} | (Not available) | Depositor |
| R_{sym} | (Not available) | Depositor |
| $\langle I/\sigma(I) \rangle$ ¹ | 1.69 (at 2.81Å) | Xtriage |
| Refinement program | PHENIX (phenix.refine: 1.8.1061) | Depositor |
| R, R_{free} | 0.250 , 0.298 0.254 , 0.259 | Depositor DCC |
| R_{free} test set | 1608 reflections (5.07%) | DCC |
| Wilson B-factor (Å ²) | 43.7 | Xtriage |
| Anisotropy | 0.216 | Xtriage |
| Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²) | 0.30 , 11.4 | EDS |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| L-test for twinning | $\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$ | Xtriage |
| Outliers | 2 of 31748 reflections (0.006%) | Xtriage |
| F_o, F_c correlation | 0.90 | EDS |
| Total number of atoms | 8580 | wwPDB-VP |
| Average B, all atoms (Å ²) | 35.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 42.00 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 2.1729e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

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: 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 | A | 0.30 | 0/2460 | 0.50 | 2/3309 (0.1%) |
| 1 | B | 0.29 | 0/2433 | 0.46 | 1/3274 (0.0%) |
| 2 | C | 0.28 | 0/1842 | 0.48 | 0/2486 |
| 2 | D | 0.24 | 0/1848 | 0.43 | 0/2494 |
| All | All | 0.28 | 0/8583 | 0.47 | 3/11563 (0.0%) |

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

| Mol | Chain | #Chirality outliers | #Planarity outliers |
|-----|-------|---------------------|---------------------|
| 1 | B | 0 | 2 |
| 2 | D | 0 | 1 |
| All | All | 0 | 3 |

There are no bond length outliers.

All (3) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|--------|------|-----------|-------|-------------|----------|
| 1 | A | 303[A] | ARG | NE-CZ-NH1 | -5.50 | 117.55 | 120.30 |
| 1 | A | 303[B] | ARG | NE-CZ-NH1 | -5.50 | 117.55 | 120.30 |
| 1 | B | 232 | GLY | N-CA-C | 5.19 | 126.08 | 113.10 |

There are no chirality outliers.

All (3) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|-----|-------|-----|------|---------|
| 1 | B | 216 | ARG | Peptide |

Continued on next page...

Continued from previous page...

| Mol | Chain | Res | Type | Group |
|-----|-------|-----|------|---------|
| 1 | B | 218 | GLU | Peptide |
| 2 | D | 243 | GLU | Peptide |

5.2 Close contacts

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | A | 2406 | 0 | 2423 | 55 | 0 |
| 1 | B | 2383 | 0 | 2394 | 57 | 0 |
| 2 | C | 1804 | 0 | 1851 | 49 | 0 |
| 2 | D | 1810 | 0 | 1855 | 37 | 0 |
| 3 | C | 1 | 0 | 0 | 0 | 0 |
| 3 | D | 1 | 0 | 0 | 0 | 0 |
| 4 | A | 55 | 0 | 0 | 2 | 0 |
| 4 | B | 44 | 0 | 0 | 7 | 0 |
| 4 | C | 26 | 0 | 0 | 2 | 0 |
| 4 | D | 50 | 0 | 0 | 2 | 0 |
| All | All | 8580 | 0 | 8523 | 190 | 0 |

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 11.

The worst 5 of 190 close contacts within the same asymmetric unit are listed below.

| Atom-1 | Atom-2 | Distance(Å) | Clash(Å) |
|-----------------|------------------|-------------|----------|
| 1:B:218:GLU:OE2 | 4:B:544:HOH:O | 1.58 | 1.18 |
| 1:B:219:ASP:OD1 | 1:B:221:SER:OG | 1.85 | 0.93 |
| 1:A:216:ARG:HB2 | 1:A:216:ARG:HH11 | 1.42 | 0.83 |
| 1:A:278:GLY:HA3 | 1:A:353:ARG:HD2 | 1.62 | 0.80 |
| 2:C:366:GLU:OE1 | 2:C:414:ARG:NH1 | 2.17 | 0.76 |

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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 | A | 287/336 (85%) | 270 (94%) | 16 (6%) | 1 (0%) | 50 | 84 |
| 1 | B | 285/336 (85%) | 272 (95%) | 12 (4%) | 1 (0%) | 43 | 80 |
| 2 | C | 221/264 (84%) | 206 (93%) | 14 (6%) | 1 (0%) | 38 | 76 |
| 2 | D | 222/264 (84%) | 209 (94%) | 9 (4%) | 4 (2%) | 13 | 39 |
| All | All | 1015/1200 (85%) | 957 (94%) | 51 (5%) | 7 (1%) | 30 | 69 |

5 of 7 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | A | 460 | SER |
| 2 | D | 244 | PRO |
| 2 | C | 321 | LYS |
| 2 | D | 478 | LYS |
| 2 | D | 243 | GLU |

5.3.2 Protein sidechains ⓘ

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 | A | 265/295 (90%) | 246 (93%) | 19 (7%) | 21 | 50 |
| 1 | B | 262/295 (89%) | 246 (94%) | 16 (6%) | 26 | 61 |
| 2 | C | 197/227 (87%) | 184 (93%) | 13 (7%) | 24 | 56 |
| 2 | D | 197/227 (87%) | 183 (93%) | 14 (7%) | 21 | 51 |
| All | All | 921/1044 (88%) | 859 (93%) | 62 (7%) | 24 | 55 |

5 of 62 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | B | 303 | ARG |
| 1 | B | 457 | GLN |
| 2 | D | 420 | LEU |
| 1 | B | 432 | THR |
| 2 | C | 234 | ARG |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 2 | C | 333 | HIS |
| 2 | C | 477 | HIS |
| 2 | C | 485 | GLN |

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

| Mol | Chain | Analysed | <RSRZ> | #RSRZ>2 | OWAB(Å ²) | Q<0.9 |
|-----|-------|-----------------|--------|---------------|-----------------------|-------|
| 1 | A | 294/336 (87%) | 0.10 | 5 (1%) 67 68 | 12, 29, 65, 86 | 0 |
| 1 | B | 293/336 (87%) | 0.23 | 11 (3%) 38 38 | 15, 31, 68, 97 | 0 |
| 2 | C | 226/264 (85%) | 0.13 | 4 (1%) 65 66 | 18, 38, 61, 80 | 0 |
| 2 | D | 227/264 (85%) | 0.00 | 5 (2%) 59 60 | 17, 31, 58, 76 | 0 |
| All | All | 1040/1200 (86%) | 0.12 | 25 (2%) 56 57 | 12, 32, 64, 97 | 0 |

The worst 5 of 25 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1 | A | 312 | ALA | 4.8 |
| 1 | B | 209 | LEU | 4.2 |
| 1 | B | 311 | THR | 3.8 |
| 2 | D | 261 | PRO | 3.8 |
| 1 | B | 319 | LEU | 3.5 |

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

| Mol | Type | Chain | Res | Atoms | RSR | LLDF | B-factors(\AA^2) | Q<0.9 |
|-----|------|-------|-----|-------|------|------|-----------------------------|-------|
| 3 | MG | D | 501 | 1/1 | 0.28 | 5.39 | 13,13,13,13 | 0 |
| 3 | MG | C | 501 | 1/1 | 0.19 | 0.65 | 10,10,10,10 | 0 |

6.5 Other polymers ⓘ

There are no such residues in this entry.