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

Aug 6, 2023 – 03:54 AM EDT

| | | |
|---|---|---|
| PDB ID | : | 1JH9 |
| Title | : | PURINE REPRESSOR MUTANT-HYPOXANTHINE-PURF OPERATOR COMPLEX |
| Authors | : | Huffman, J.L.; Lu, F.; Zalkin, H.; Brennan, R.G. |
| Deposited on | : | 2001-06-27 |
| Resolution | : | 2.55 Å(reported) |

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
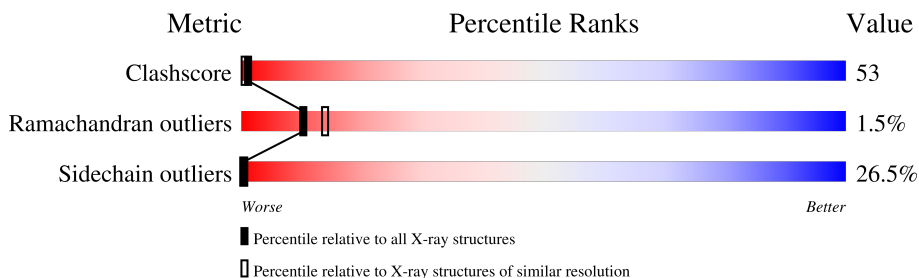http://www.wwpdb.org/validation/2017/FAQs#types.

---

# 1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.55 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 | | 53 |
| Ramachandran outliers | | 1.5% |
| Sidechain outliers | | 26.5% |

*Worse* — *Better*

■ 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 | 141614 | 1332 (2.56-2.52) |
| Ramachandran outliers | 138981 | 1315 (2.56-2.52) |
| Sidechain outliers | 138945 | 1315 (2.56-2.52) |

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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%

Note EDS was not executed.

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | B | 17 | 24% 71% 6% |
| 2 | A | 340 | 28% 54% 18% • |

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol | Type | Chain | Res | Chirality | Geometry | Clashes | Electron density |
|---|---|---|---|---|---|---|---|
| 3 | PO4 | A | 799 | - | - | X | - |
| 3 | PO4 | A | 800 | - | X | - | - |

## 2  Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 3136 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 DNA chain called 5'-D(*AP*AP*CP*GP*AP*AP*AP*AP*CP*GP*TP*T P*TP*TP*CP*GP*T)-3'.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 1 | B | 17 | Total 346 | C 167 | N 64 | O 99 | P 16 | 0 | 0 | 0 |

- Molecule 2 is a protein called PURINE NUCLEOTIDE SYNTHESIS REPRESSOR.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 2 | A | 339 | Total 2654 | C 1669 | N 472 | O 494 | S 19 | 0 | 0 | 0 |

There is a discrepancy between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| A | 147 | ARG | TRP | engineered mutation | UNP P0ACP7 |

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: $O_4P$).
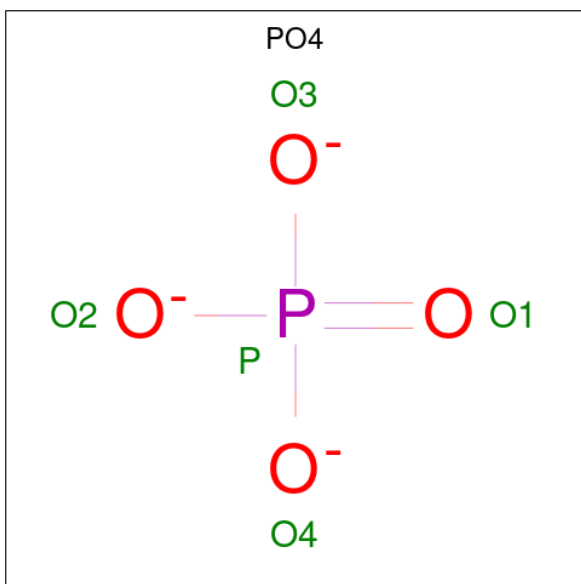
| Mol | Chain | Residues | Atoms | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|
| 3 | A | 1 | Total 5 | O 4 | P 1 | 0 | 0 |
| 3 | A | 1 | Total 5 | O 4 | P 1 | 0 | 0 |

- Molecule 4 is HYPOXANTHINE (three-letter code: HPA) (formula: $C_5H_4N_4O$).

HPA

O6

C6

N1 N7

C5

C2 C8

C4

N3 N9
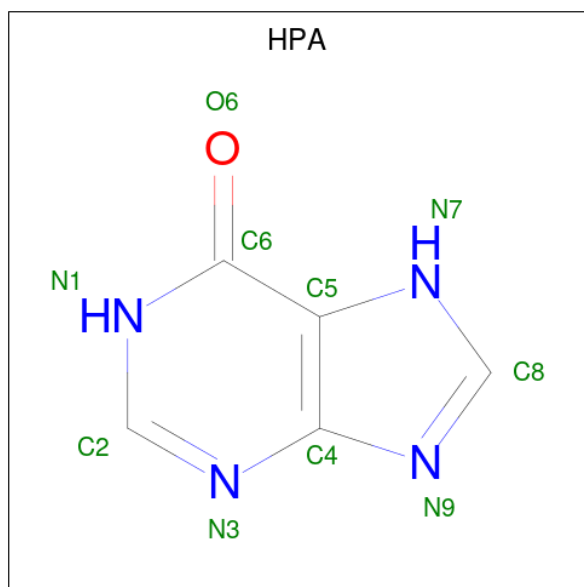
| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|
| 4 | A | 1 | Total 10 | C 5 | N 4 | O 1 | 0 | 0 |

- Molecule 5 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 5 | B | 9 | Total 9 | O 9 | 0 | 0 |
| 5 | A | 107 | Total 107 | O 107 | 0 | 0 |

W O R L D W I D E

PDB
PROTEIN DATA BANK

# 3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.
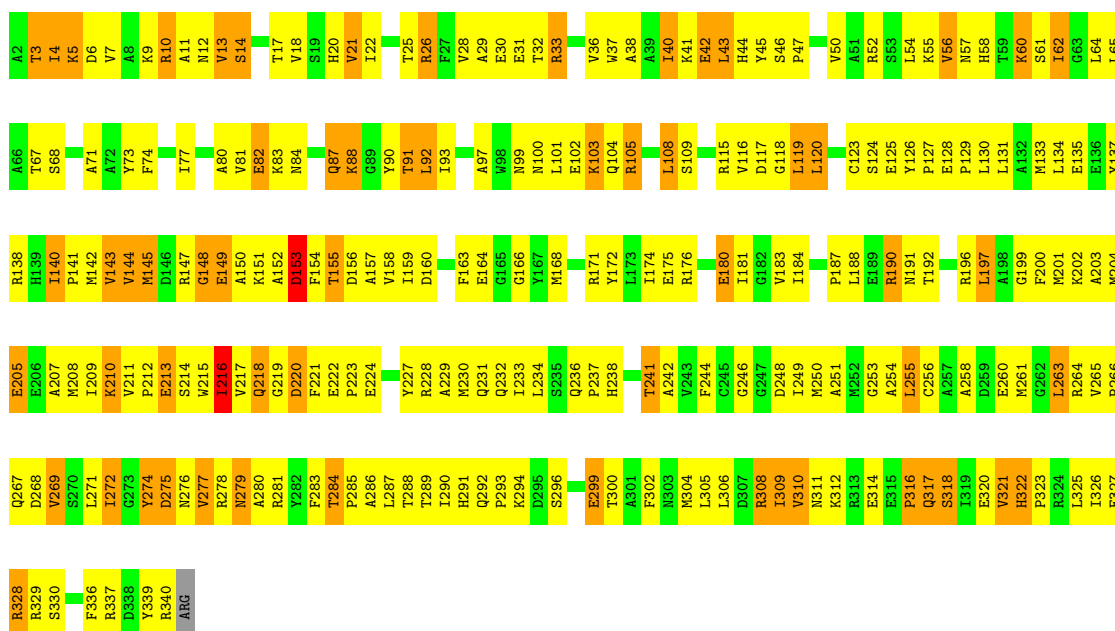
Note EDS was not executed.

• Molecule 1: 5'-D(*AP*AP*CP*GP*AP*AP*AP*AP*CP*GP*TP*TP*TP*TP*CP*GP*T)-3'

Chain B:

24%     71%     6%

A699 A700 C701 G702 A703 A704 C707 G708 T709 T710 T711 T712 C713 G714 T715

• Molecule 2: PURINE NUCLEOTIDE SYNTHESIS REPRESSOR

Chain A:

28%     54%     18%

A2 T3 I4 K5 D6 V7 A8 K9 R10 A11 N12 V13 S14 T17 V18 S19 H20 V21 I22 T25 R26 F27 V28 A29 E30 E31 T32 R33 V36 W37 A38 A39 I40 K41 E42 L43 H44 Y45 S46 P47 V50 A51 R52 S53 L54 K55 V56 N57 H58 T59 K60 S61 I62 G63 L64 L65

A66 T67 S68 A71 A72 Y73 F74 I77 A80 V81 E82 K83 N84 Q87 K88 G89 Y90 T91 L92 I93 A97 W98 N99 N100 L101 E102 K103 Q104 R105 L108 S109 R115 V116 D117 G118 L119 L120 C123 S124 E125 Y126 P127 E128 P129 L130 L131 A132 M133 L134 E135 E136 Y137

R138 H139 I140 P141 M142 V143 V144 M145 D146 R147 G148 E149 A150 K151 A152 D153 F154 T155 D156 A157 V158 I159 D160 F163 E164 G165 G166 Y167 M168 R171 Y172 L173 I174 E175 R176 E180 I181 G182 V183 I184 P187 L188 E189 R190 N191 T192 R196 L197 A198 G199 F200 K202 A203 M204

E205 E206 A207 M208 I209 K210 V211 P212 E213 S214 W215 I216 V217 Q218 G219 D220 F221 E222 P223 E224 Y227 R228 A229 M230 Q231 Q232 I233 L234 S235 Q236 P237 H238 T241 A242 V243 F244 C245 G246 Q247 D248 I249 M250 A251 M252 G253 A254 L255 C256 A257 A258 D259 E260 M261 G262 L263 R264 V265 P266

Q267 D268 V269 S270 L271 I272 G273 Y274 D275 N276 V277 R278 N279 A280 R281 Y282 F283 T284 P285 A286 L287 T288 T289 I290 H291 Q292 P293 K294 D295 S296 E299 T300 A301 F302 N303 M304 L305 L306 D307 R308 I309 V310 N311 K312 R313 E314 E315 P316 Q317 S318 I319 E320 V321 H322 P323 R324 L325 I326 E327

R328 R329 S330 F336 R337 D338 Y339 R340 ARG

# 4 Data and refinement statistics ⓘ

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

| Property | Value | Source |
|---|---|---|
| Space group | C 2 2 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 175.38Å   92.61Å   80.64Å<br>90.00°   90.00°   90.00° | Depositor |
| Resolution (Å) | 10.00   –   2.55 | Depositor |
| % Data completeness<br>(in resolution range) | 96.0 (10.00-2.55) | Depositor |
| $R_{merge}$ | 0.11 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| Refinement program | TNT | Depositor |
| R, $R_{free}$ | 0.219   ,   (Not available) | Depositor |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 3136 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 57.0 | wwPDB-VP |

# 5 Model quality ⓘ

## 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, HPA

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 | 0.71 | 0/388 | 0.98 | 1/597 (0.2%) |
| 2 | A | 1.02 | 0/2706 | 1.09 | 8/3658 (0.2%) |
| All | All | 0.99 | 0/3094 | 1.08 | 9/4255 (0.2%) |

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|-----|----------------|-------------|
| 2 | A | 216 | ILE | CG1-CB-CG2 | -8.54 | 92.61 | 111.40 |
| 2 | A | 340 | ARG | NE-CZ-NH2 | 7.42 | 124.01 | 120.30 |
| 2 | A | 220 | ASP | CB-CG-OD1 | 6.42 | 124.08 | 118.30 |
| 2 | A | 180 | GLU | N-CA-C | -5.83 | 95.25 | 111.00 |
| 2 | A | 322 | HIS | N-CA-C | 5.62 | 126.19 | 111.00 |

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-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 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 | 346 | 0 | 194 | 13 | 0 |
| 2 | A | 2654 | 0 | 2644 | 299 | 0 |
| 3 | A | 10 | 0 | 0 | 2 | 0 |
| 4 | A | 10 | 0 | 4 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 5 | A | 107 | 0 | 0 | 9 | 1 |
| 5 | B | 9 | 0 | 0 | 0 | 0 |
| All | All | 3136 | 0 | 2842 | 310 | 1 |

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 53.

The worst 5 of 310 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:A:210:LYS:H | 2:A:210:LYS:HD3 | 1.18 | 1.04 |
| 2:A:223:PRO:HD3 | 2:A:249:ILE:HG22 | 1.38 | 1.03 |
| 2:A:126:TYR:HB3 | 2:A:131:LEU:HD21 | 1.44 | 0.98 |
| 2:A:183:VAL:CG1 | 2:A:216:ILE:HG22 | 2.03 | 0.89 |
| 2:A:101:LEU:HA | 2:A:104:GLN:NE2 | 1.88 | 0.88 |

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|--------|--------|--------------------------|-------------------|
| 5:A:847:HOH:O | 5:A:847:HOH:O[4_555] | 2.12 | 0.08 |

## 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 | |
|-----|-------|----------|----------|---------|----------|-------------|---|
| 2 | A | 337/340 (99%) | 296 (88%) | 36 (11%) | 5 (2%) | 10 | 14 |

All (5) Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 2 | A | 246 | GLY |
| 2 | A | 275 | ASP |
| 2 | A | 153 | ASP |
| 2 | A | 148 | GLY |
| 2 | A | 309 | ILE |

### 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 | |
|-----|-------|----------|-----------|----------|---|---|
| 2 | A | 279/280 (100%) | 205 (74%) | 74 (26%) | 0 | 0 |

5 of 74 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 2 | A | 263 | LEU |
| 2 | A | 318 | SER |
| 2 | A | 272 | ILE |
| 2 | A | 296 | SER |
| 2 | A | 92 | LEU |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 2 | A | 279 | ASN |
| 2 | A | 291 | HIS |
| 2 | A | 322 | HIS |
| 2 | A | 104 | GLN |
| 2 | A | 96 | ASN |

### 5.3.3  RNA  ⓘ

There are no RNA molecules 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

3 ligands are modelled in this entry.

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 | HPA | A | 599 | - | 8,11,11 | 1.84 | 2 (25%) | 5,15,15 | 2.96 | 2 (40%) |
| 3 | PO4 | A | 799 | - | 4,4,4 | 2.34 | 2 (50%) | 6,6,6 | 1.55 | 1 (16%) |
| 3 | PO4 | A | 800 | - | 4,4,4 | 2.61 | 4 (100%) | 6,6,6 | 0.73 | 0 |

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 | HPA | A | 599 | - | - | - | 0/2/2/2 |

The worst 5 of 8 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|------|-------------|----------|
| 4 | A | 599 | HPA | C6-N1 | 3.93 | 1.39 | 1.33 |
| 3 | A | 799 | PO4 | P-O2 | -3.56 | 1.43 | 1.54 |
| 3 | A | 799 | PO4 | P-O3 | -3.02 | 1.45 | 1.54 |
| 3 | A | 800 | PO4 | P-O2 | -2.97 | 1.45 | 1.54 |
| 3 | A | 800 | PO4 | P-O1 | 2.69 | 1.57 | 1.50 |

All (3) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|---|----------------|-------------|
| 4 | A | 599 | HPA | C2-N1-C6 | 5.47 | 125.04 | 115.88 |
| 4 | A | 599 | HPA | N3-C2-N1 | -3.20 | 123.67 | 128.68 |
| 3 | A | 799 | PO4 | O4-P-O3 | 2.57 | 116.21 | 107.97 |

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 3 | A | 799 | PO4 | 2 | 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 ⓘ

## 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

## 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

## 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

## 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.