



Full wwPDB NMR Structure Validation Report ⓘ

Mar 6, 2022 – 01:10 PM EST

PDB ID : 2KHW
Title : Solution Structure of the human Polymerase iota UBM2-Ubiquitin Complex
Authors : Bomar, M.G.; D'Souza, S.; Bienko, M.; Dikic, I.; Walker, G.
Deposited on : 2009-04-13

This is a Full wwPDB NMR Structure Validation 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/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
ShiftChecker : 2.27
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.27

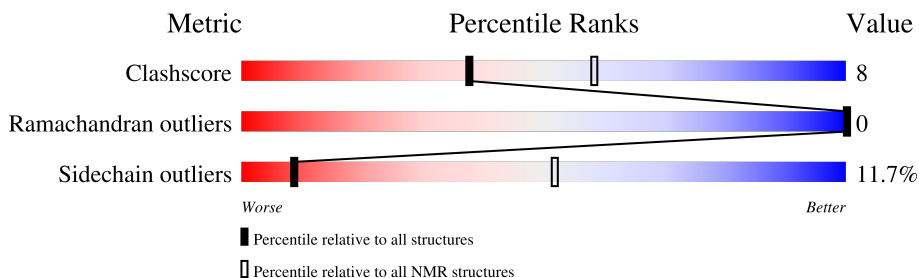
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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) | NMR archive (#Entries) |
|-----------------------|--------------------------|------------------------|
| Clashscore | 158937 | 12864 |
| Ramachandran outliers | 154571 | 11451 |
| Sidechain outliers | 154315 | 11428 |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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 $\leq 5\%$

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | A | 108 | |
| 2 | B | 79 | |

2 Ensemble composition and analysis i

This entry contains 25 models. Model 4 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

| Well-defined (core) protein residues | | | |
|--------------------------------------|--------------------------------|-------------------|--------------|
| Well-defined core | Residue range (total) | Backbone RMSD (Å) | Medoid model |
| 1 | A:64-A:88, B:201-B:272 (97) | 0.12 | 4 |

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 6 clusters and 1 single-model cluster was found.

| Cluster number | Models |
|-----------------------|-----------------------------|
| 1 | 6, 7, 9, 12, 15, 17, 19, 25 |
| 2 | 1, 8, 16, 18 |
| 3 | 2, 13, 21, 24 |
| 4 | 4, 11, 14 |
| 5 | 3, 10, 20 |
| 6 | 22, 23 |
| Single-model clusters | 5 |

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1743 atoms, of which 885 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Immunoglobulin G-binding protein G, DNA polymerase iota.

| Mol | Chain | Residues | Atoms | | | | | Trace |
|-----|-------|----------|-------|-----|-----|----|----|-------|
| | | | Total | C | H | N | O | |
| 1 | A | 32 | 528 | 174 | 263 | 41 | 50 | 0 |

There are 14 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|----------------|------------|
| A | 1 | MET | - | expression tag | UNP P19909 |
| A | 2 | GLN | - | expression tag | UNP P19909 |
| A | 57 | GLY | - | SEE REMARK 999 | UNP Q9UNA4 |
| A | 58 | SER | - | SEE REMARK 999 | UNP Q9UNA4 |
| A | 59 | ASP | - | SEE REMARK 999 | UNP Q9UNA4 |
| A | 60 | GLU | - | SEE REMARK 999 | UNP Q9UNA4 |
| A | 101 | LEU | - | expression tag | UNP Q9UNA4 |
| A | 102 | GLU | - | expression tag | UNP Q9UNA4 |
| A | 103 | HIS | - | expression tag | UNP Q9UNA4 |
| A | 104 | HIS | - | expression tag | UNP Q9UNA4 |
| A | 105 | HIS | - | expression tag | UNP Q9UNA4 |
| A | 106 | HIS | - | expression tag | UNP Q9UNA4 |
| A | 107 | HIS | - | expression tag | UNP Q9UNA4 |
| A | 108 | HIS | - | expression tag | UNP Q9UNA4 |

- Molecule 2 is a protein called Ubiquitin.

| Mol | Chain | Residues | Atoms | | | | | Trace | |
|-----|-------|----------|-------|-----|-----|-----|-----|-------|---|
| | | | Total | C | H | N | O | | S |
| 2 | B | 74 | 1215 | 374 | 622 | 103 | 115 | 1 | 0 |

There are 3 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|----------------|------------|
| B | 198 | GLY | - | expression tag | UNP P62988 |
| B | 199 | SER | - | expression tag | UNP P62988 |
| B | 200 | HIS | - | expression tag | UNP P62988 |

MET GLN TYR LYS LEU ILE LEU ASN GLY LYS THR LEU LYS GLY THR THR ALA VAL ASP PHE ALA ALA THR ALA GLU VAL PHE HIS LYS THR TYR ALA ASN ASP ASN GLY VAL ASP GLY TRP THR TYR ASP ASP ALA THR LYS THR PHE THR VAL THR GLY SER ASP GLU



- Molecule 2: Ubiquitin



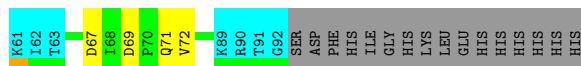
GLY SER HIS M201 Q202 K206 T207 L208 T209 S220 E224 D232 D239 Q240 Q241 K248 D252 L256 Y259 N260 K263 T266 L267 H268 L269 L273 R274 GLY

4.2.20 Score per residue for model 20

- Molecule 1: Immunoglobulin G-binding protein G, DNA polymerase iota



MET GLN TYR LYS LEU ILE LEU ASN GLY LYS THR LEU LYS GLY THR THR ALA VAL ASP PHE ALA ALA THR ALA GLU VAL PHE HIS LYS THR TYR ALA ASN ASP ASN GLY VAL ASP GLY TRP THR TYR ASP ASP ALA THR LYS THR PHE THR VAL THR GLY SER ASP GLU



- Molecule 2: Ubiquitin



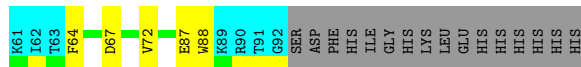
GLY SER HIS M201 Q202 K206 T207 L208 T209 L215 E224 K227 A228 K229 Q240 F245 A246 D252 L256 S257 D258 Y259 N260 K263 T266 L267 L273 R274 GLY

4.2.21 Score per residue for model 21

- Molecule 1: Immunoglobulin G-binding protein G, DNA polymerase iota



MET GLN TYR LYS LEU ILE LEU ASN GLY LYS THR LEU LYS GLY THR THR ALA VAL ASP PHE ALA ALA THR ALA GLU VAL PHE HIS LYS THR TYR ALA ASN ASP ASN GLY VAL ASP GLY TRP THR TYR ASP ASP ALA THR LYS THR PHE THR VAL THR GLY SER ASP GLU



- Molecule 2: Ubiquitin



5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing*.

Of the 40 calculated structures, 25 were deposited, based on the following criterion: *structures with the lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

| Software name | Classification | Version |
|---------------|--------------------|---------|
| X-PLOR NIH | refinement | 2.19 |
| CYANA | structure solution | |

No chemical shift data was provided.

6 Model quality

6.1 Standard geometry

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 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 each 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 averaged over the ensemble.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes |
|-----|-------|-------|----------|----------|---------|
| 1 | A | 210 | 196 | 196 | 4±1 |
| 2 | B | 574 | 598 | 596 | 11±3 |
| All | All | 19600 | 19850 | 19800 | 318 |

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 8.

All unique clashes are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models | |
|------------------|------------------|----------|-------------|--------|-------|
| | | | | Worst | Total |
| 1:A:67:ASP:O | 2:B:209:THR:HG22 | 0.83 | 1.73 | 13 | 25 |
| 2:B:206:LYS:NZ | 2:B:266:THR:HG21 | 0.76 | 1.95 | 19 | 3 |
| 2:B:266:THR:C | 2:B:267:LEU:HD12 | 0.73 | 2.04 | 16 | 25 |
| 2:B:259:TYR:O | 2:B:260:ASN:ND2 | 0.66 | 2.29 | 23 | 2 |
| 2:B:216:GLU:O | 2:B:229:LYS:NZ | 0.65 | 2.28 | 7 | 2 |
| 1:A:72:VAL:HG11 | 2:B:207:THR:O | 0.63 | 1.94 | 6 | 22 |
| 2:B:260:ASN:ND2 | 2:B:260:ASN:O | 0.63 | 2.32 | 6 | 1 |
| 2:B:213:ILE:HD11 | 2:B:234:GLU:OE2 | 0.62 | 1.95 | 23 | 3 |
| 2:B:215:LEU:C | 2:B:229:LYS:HZ3 | 0.60 | 1.99 | 7 | 2 |
| 2:B:251:GLU:OE1 | 2:B:252:ASP:N | 0.59 | 2.36 | 15 | 2 |
| 2:B:263:LYS:H | 2:B:263:LYS:NZ | 0.59 | 1.96 | 9 | 2 |
| 2:B:222:THR:OG1 | 2:B:225:ASN:ND2 | 0.59 | 2.36 | 4 | 6 |
| 2:B:206:LYS:HZ3 | 2:B:266:THR:HG21 | 0.58 | 1.55 | 19 | 2 |

Continued on next page...

Continued from previous page...

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models | |
|------------------|------------------|----------|-------------|--------|-------|
| | | | | Worst | Total |
| 2:B:213:ILE:HD11 | 2:B:234:GLU:CD | 0.58 | 2.19 | 11 | 3 |
| 1:A:69:ASP:OD1 | 1:A:71:GLN:N | 0.57 | 2.37 | 24 | 15 |
| 2:B:252:ASP:N | 2:B:252:ASP:OD1 | 0.57 | 2.37 | 21 | 7 |
| 2:B:260:ASN:O | 2:B:260:ASN:CG | 0.56 | 2.44 | 6 | 2 |
| 2:B:267:LEU:HD12 | 2:B:267:LEU:N | 0.56 | 2.16 | 8 | 22 |
| 2:B:206:LYS:NZ | 2:B:206:LYS:CB | 0.55 | 2.69 | 16 | 4 |
| 2:B:217:VAL:HG13 | 2:B:229:LYS:NZ | 0.55 | 2.16 | 7 | 2 |
| 2:B:216:GLU:C | 2:B:229:LYS:HZ1 | 0.54 | 2.04 | 3 | 2 |
| 2:B:260:ASN:ND2 | 2:B:260:ASN:C | 0.54 | 2.60 | 6 | 1 |
| 2:B:263:LYS:H | 2:B:263:LYS:HZ3 | 0.54 | 1.44 | 9 | 1 |
| 2:B:251:GLU:OE1 | 2:B:253:GLY:N | 0.54 | 2.40 | 17 | 1 |
| 2:B:237:PRO:C | 2:B:241:GLN:HE21 | 0.53 | 2.06 | 1 | 1 |
| 2:B:260:ASN:OD1 | 2:B:260:ASN:O | 0.53 | 2.26 | 23 | 1 |
| 2:B:248:LYS:NZ | 2:B:248:LYS:CB | 0.52 | 2.72 | 9 | 2 |
| 2:B:206:LYS:HZ1 | 2:B:266:THR:HG21 | 0.52 | 1.64 | 2 | 2 |
| 2:B:215:LEU:C | 2:B:229:LYS:NZ | 0.52 | 2.62 | 7 | 2 |
| 1:A:87:GLU:OE1 | 1:A:88:TRP:CD1 | 0.52 | 2.62 | 10 | 2 |
| 2:B:245:PHE:O | 2:B:246:ALA:HB3 | 0.52 | 2.05 | 11 | 15 |
| 1:A:67:ASP:N | 1:A:67:ASP:OD1 | 0.51 | 2.43 | 5 | 8 |
| 2:B:227:LYS:NZ | 2:B:252:ASP:OD1 | 0.51 | 2.43 | 25 | 1 |
| 2:B:239:ASP:OD1 | 2:B:240:GLN:N | 0.51 | 2.44 | 4 | 2 |
| 2:B:264:GLU:N | 2:B:264:GLU:CD | 0.50 | 2.65 | 23 | 1 |
| 2:B:215:LEU:HD13 | 2:B:229:LYS:CB | 0.49 | 2.37 | 1 | 9 |
| 2:B:256:LEU:N | 2:B:256:LEU:HD23 | 0.49 | 2.22 | 2 | 21 |
| 2:B:259:TYR:C | 2:B:260:ASN:HD22 | 0.49 | 2.11 | 19 | 1 |
| 2:B:263:LYS:N | 2:B:263:LYS:CD | 0.49 | 2.76 | 2 | 1 |
| 2:B:215:LEU:CB | 2:B:229:LYS:NZ | 0.48 | 2.76 | 3 | 2 |
| 1:A:70:PRO:O | 1:A:74:TYR:CD2 | 0.48 | 2.67 | 8 | 1 |
| 2:B:241:GLN:HB2 | 2:B:269:LEU:HD11 | 0.48 | 1.85 | 23 | 1 |
| 2:B:208:LEU:HD11 | 2:B:271:LEU:H | 0.47 | 1.69 | 24 | 2 |
| 2:B:255:THR:OG1 | 2:B:258:ASP:OD2 | 0.47 | 2.32 | 8 | 1 |
| 2:B:262:GLN:O | 2:B:265:SER:OG | 0.47 | 2.31 | 21 | 2 |
| 2:B:264:GLU:N | 2:B:264:GLU:OE2 | 0.47 | 2.48 | 23 | 1 |
| 2:B:248:LYS:CB | 2:B:248:LYS:NZ | 0.46 | 2.79 | 7 | 1 |
| 2:B:206:LYS:CE | 2:B:268:HIS:NE2 | 0.46 | 2.79 | 2 | 1 |
| 2:B:215:LEU:HD13 | 2:B:229:LYS:HB2 | 0.46 | 1.88 | 24 | 9 |
| 2:B:203:ILE:HG22 | 2:B:265:SER:OG | 0.46 | 2.11 | 4 | 4 |
| 2:B:237:PRO:C | 2:B:241:GLN:NE2 | 0.46 | 2.69 | 1 | 1 |
| 2:B:252:ASP:OD1 | 2:B:252:ASP:N | 0.45 | 2.48 | 10 | 2 |
| 2:B:215:LEU:HB3 | 2:B:229:LYS:NZ | 0.45 | 2.26 | 7 | 2 |
| 2:B:206:LYS:NZ | 2:B:266:THR:CG2 | 0.45 | 2.77 | 2 | 2 |

Continued on next page...

Continued from previous page...

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models | |
|------------------|------------------|----------|-------------|--------|-------|
| | | | | Worst | Total |
| 2:B:204:PHE:CD1 | 2:B:204:PHE:N | 0.45 | 2.81 | 22 | 4 |
| 2:B:215:LEU:CB | 2:B:229:LYS:HZ3 | 0.45 | 2.24 | 3 | 1 |
| 1:A:69:ASP:OD1 | 1:A:69:ASP:C | 0.44 | 2.55 | 13 | 5 |
| 2:B:230:ILE:O | 2:B:234:GLU:N | 0.44 | 2.49 | 2 | 2 |
| 2:B:230:ILE:HD12 | 2:B:230:ILE:H | 0.44 | 1.72 | 11 | 4 |
| 2:B:224:GLU:OE1 | 2:B:224:GLU:O | 0.44 | 2.34 | 10 | 1 |
| 1:A:69:ASP:C | 1:A:69:ASP:OD1 | 0.44 | 2.56 | 18 | 1 |
| 2:B:250:LEU:HD11 | 2:B:267:LEU:HD23 | 0.44 | 1.87 | 8 | 3 |
| 2:B:224:GLU:O | 2:B:224:GLU:CD | 0.43 | 2.57 | 23 | 1 |
| 2:B:255:THR:OG1 | 2:B:258:ASP:OD1 | 0.43 | 2.36 | 10 | 1 |
| 2:B:224:GLU:CD | 2:B:224:GLU:C | 0.43 | 2.77 | 23 | 5 |
| 1:A:73:PHE:CE1 | 1:A:88:TRP:CH2 | 0.43 | 3.06 | 19 | 6 |
| 2:B:266:THR:O | 2:B:267:LEU:HD12 | 0.43 | 2.13 | 19 | 2 |
| 2:B:213:ILE:CD1 | 2:B:234:GLU:CG | 0.43 | 2.96 | 11 | 2 |
| 2:B:260:ASN:OD1 | 2:B:260:ASN:N | 0.43 | 2.51 | 25 | 1 |
| 2:B:239:ASP:C | 2:B:239:ASP:OD1 | 0.43 | 2.57 | 1 | 2 |
| 2:B:263:LYS:HD2 | 2:B:263:LYS:H | 0.43 | 1.73 | 2 | 4 |
| 2:B:263:LYS:CD | 2:B:263:LYS:H | 0.43 | 2.26 | 2 | 1 |
| 2:B:217:VAL:CG1 | 2:B:229:LYS:NZ | 0.42 | 2.83 | 3 | 1 |
| 1:A:77:PRO:O | 1:A:78:GLU:C | 0.42 | 2.57 | 23 | 1 |
| 2:B:224:GLU:CD | 2:B:225:ASN:OD1 | 0.42 | 2.58 | 18 | 1 |
| 2:B:224:GLU:CG | 2:B:225:ASN:N | 0.42 | 2.83 | 17 | 1 |
| 2:B:241:GLN:HG3 | 2:B:269:LEU:HD11 | 0.41 | 1.93 | 19 | 1 |
| 1:A:66:SER:C | 1:A:67:ASP:OD1 | 0.41 | 2.58 | 11 | 2 |
| 2:B:239:ASP:OD1 | 2:B:239:ASP:C | 0.41 | 2.58 | 4 | 2 |
| 2:B:217:VAL:HG13 | 2:B:229:LYS:HZ2 | 0.41 | 1.76 | 7 | 1 |
| 2:B:254:ARG:NH1 | 2:B:259:TYR:OH | 0.41 | 2.52 | 13 | 1 |
| 2:B:238:PRO:N | 2:B:241:GLN:HE21 | 0.40 | 2.14 | 1 | 1 |
| 2:B:260:ASN:N | 2:B:260:ASN:ND2 | 0.40 | 2.70 | 19 | 1 |
| 2:B:245:PHE:O | 2:B:246:ALA:CB | 0.40 | 2.69 | 11 | 1 |

6.3 Torsion angles [i](#)

6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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 | 25/108 (23%) | 24±0 (94±2%) | 1±0 (6±2%) | 0±0 (0±0%) | 100 | 100 |
| 2 | B | 71/79 (90%) | 66±1 (93±1%) | 5±1 (7±1%) | 0±0 (0±0%) | 100 | 100 |
| All | All | 2400/4675 (51%) | 2245 (94%) | 155 (6%) | 0 (0%) | 100 | 100 |

There are no Ramachandran outliers.

6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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 | 23/93 (25%) | 22±1 (95±4%) | 1±1 (5±4%) | 30 | 79 |
| 2 | B | 66/70 (94%) | 57±2 (86±3%) | 9±2 (14±3%) | 6 | 46 |
| All | All | 2225/4075 (55%) | 1964 (88%) | 261 (12%) | 9 | 52 |

All 26 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 2 | B | 263 | LYS | 25 |
| 2 | B | 240 | GLN | 24 |
| 2 | B | 202 | GLN | 23 |
| 2 | B | 241 | GLN | 21 |
| 2 | B | 252 | ASP | 18 |
| 2 | B | 260 | ASN | 18 |
| 2 | B | 206 | LYS | 15 |
| 1 | A | 66 | SER | 14 |
| 2 | B | 224 | GLU | 14 |
| 2 | B | 232 | ASP | 11 |
| 2 | B | 258 | ASP | 11 |
| 2 | B | 221 | ASP | 11 |
| 2 | B | 248 | LYS | 10 |
| 2 | B | 220 | SER | 8 |
| 1 | A | 67 | ASP | 6 |
| 2 | B | 251 | GLU | 6 |
| 1 | A | 78 | GLU | 5 |
| 2 | B | 227 | LYS | 5 |
| 2 | B | 254 | ARG | 4 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 2 | B | 262 | GLN | 3 |
| 1 | A | 71 | GLN | 2 |
| 2 | B | 212 | THR | 2 |
| 2 | B | 265 | SER | 2 |
| 2 | B | 218 | GLU | 1 |
| 2 | B | 216 | GLU | 1 |
| 2 | B | 264 | GLU | 1 |

6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation

No chemical shift data were provided