



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 20, 2024 – 08:35 pm GMT

PDB ID : 7AGB  
Title : Protease Sapp1p from *Candida parapsilosis* in complex with KB70  
Authors : Dostal, J.; Heidingsfeld, O.; Brynda, J.  
Deposited on : 2020-09-22  
Resolution : 1.70 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](mailto: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 ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

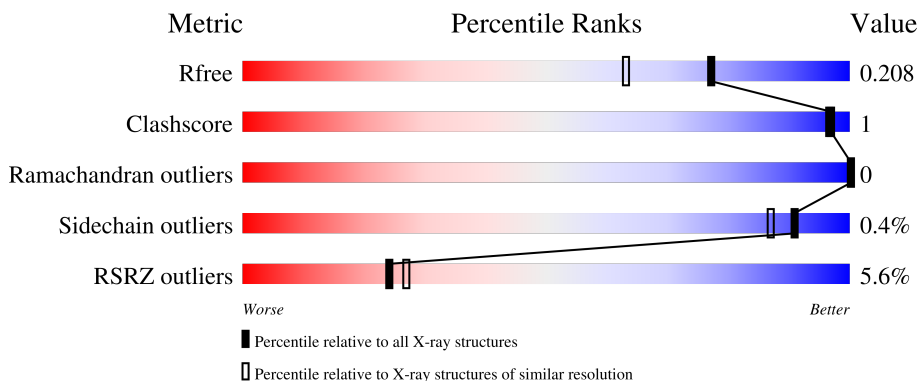
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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))
$R_{free}$	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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  $\geq 3$ , 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  $\leq 5\%$ . 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.

Mol	Chain	Length	Quality of chain
1	A	339	 4% 98%
1	B	339	 3% 97%
1	C	339	 10% 99%
1	D	339	 6% 98%
2	I	7	 14% 57% 14% 29%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	J	7	 71% 14% 14%
2	K	7	 14% 71% 14% 14%
2	L	7	 71% 14% 14%

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 11451 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 Candidapepsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	339	Total 2549	C 1589	N 414	O 542	S 4	1	9	0
1	B	339	Total 2592	C 1617	N 422	O 549	S 4	0	15	0
1	C	339	Total 2539	C 1580	N 417	O 538	S 4	0	4	0
1	D	339	Total 2566	C 1598	N 418	O 546	S 4	0	12	0

- Molecule 2 is a protein called KB70.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	I	7	Total 55	C 40	N 6	O 9	0	0	1
2	J	7	Total 55	C 40	N 6	O 9	0	0	1
2	K	7	Total 55	C 40	N 6	O 9	0	0	1
2	L	7	Total 55	C 40	N 6	O 9	0	0	1

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Na 1	0	0
3	B	1	Total 1	Na 1	0	0
3	C	1	Total 1	Na 1	0	0
3	D	3	Total 3	Na 3	0	0

- Molecule 4 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



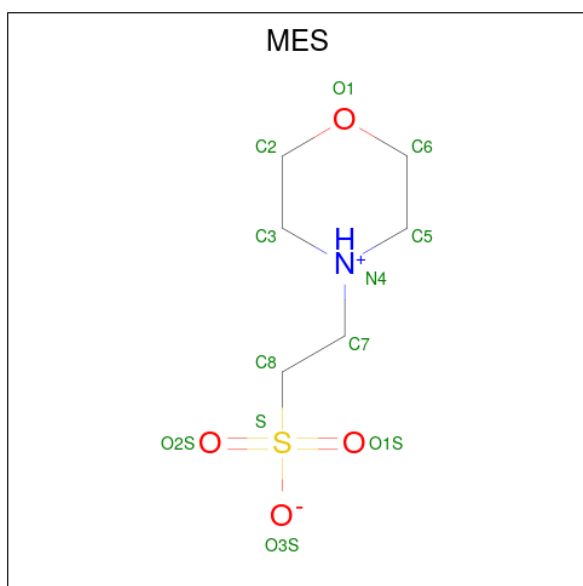
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 10 6 4	0	0
4	A	1	Total C O 7 4 3	0	0
4	A	1	Total C O 10 6 4	0	0
4	B	1	Total C O 10 6 4	0	0
4	B	1	Total C O 10 6 4	0	0
4	C	1	Total C O 9 6 3	0	0
4	C	1	Total C O 10 6 4	0	0
4	C	1	Total C O 10 6 4	0	0
4	D	1	Total C O 10 6 4	0	0
4	D	1	Total C O 10 6 4	0	0
4	D	1	Total C O 10 6 4	0	0

- Molecule 5 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula:  $C_{10}H_{22}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			15	9	6		
5	D	1	Total	C	O	0	0
			13	8	5		

- Molecule 6 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



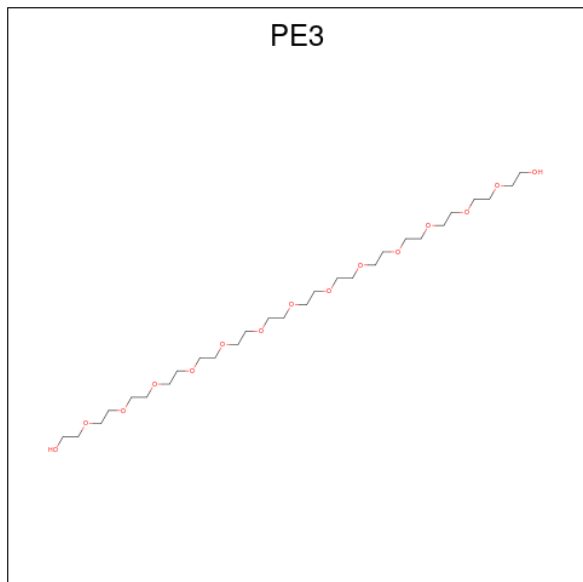
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
6	B	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
6	C	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
6	C	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
6	D	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 7 is 3,6,9,12,15,18,21,24,27,30,33,36,39-TRIDECAXAHENTETRACONTANE-1,41-DIOL (three-letter code: PE3) (formula:  $C_{28}H_{58}O_{15}$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			16	10	6		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	191	Total	O	0	1
			192	192		
8	B	233	Total	O	0	1
			234	234		
8	C	134	Total	O	0	0
			134	134		

Continued on next page...

*Continued from previous page...*

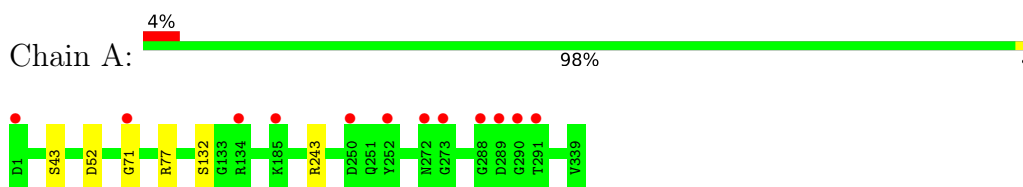
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	D	189	Total O 190 190	0	1
8	I	1	Total O 1 1	0	0
8	J	2	Total O 2 2	0	0
8	K	2	Total O 2 2	0	0
8	L	2	Total O 2 2	0	0



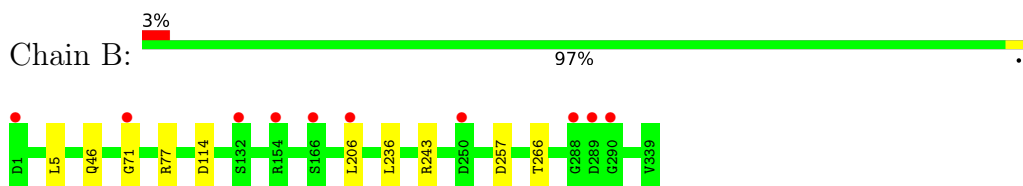
### 3 Residue-property plots [i](#)

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 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.

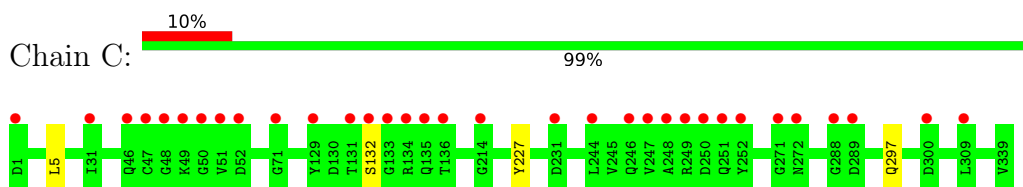
- Molecule 1: Candidapepsin



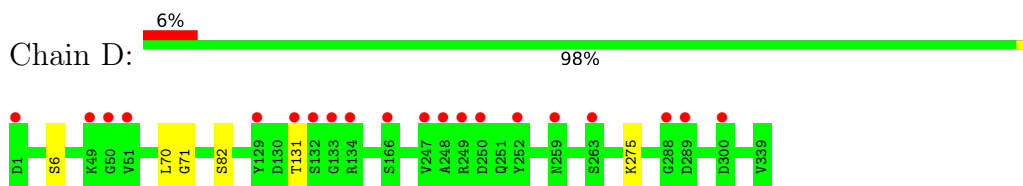
- Molecule 1: Candidapepsin



- Molecule 1: Candidapepsin



- Molecule 1: Candidapepsin



- Molecule 2: KB70




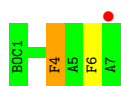
- Molecule 2: KB70

Chain J:  71% 14% 14%



● Molecule 2: KB70

Chain K:  14% 71% 14% 14%



● Molecule 2: KB70

Chain L:  71% 14% 14%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	172.02Å 172.02Å 252.69Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.58 – 1.70 34.58 – 1.70	Depositor EDS
% Data completeness (in resolution range)	97.5 (34.58-1.70) 97.5 (34.58-1.70)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.54 (at 1.70Å)	Xtrriage
Refinement program	REFMAC 5.8.0103	Depositor
R, $R_{free}$	0.179 , 0.197 0.193 , 0.208	Depositor DCC
$R_{free}$ test set	11694 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.8	Xtrriage
Anisotropy	0.496	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 48.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11451	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.54 % 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 3.6310e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 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: BOC, PGE, MES, 1PE, NA, PE3, PSA

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	0.75	4/2618 (0.2%)	0.81	3/3564 (0.1%)
1	B	0.69	0/2665	0.83	5/3628 (0.1%)
1	C	0.61	0/2594	0.72	0/3532
1	D	0.70	0/2633	0.79	0/3586
2	I	0.91	0/17	0.83	0/21
2	J	0.85	0/17	0.78	0/21
2	K	1.10	0/17	0.76	0/21
2	L	0.91	0/17	0.67	0/21
All	All	0.69	4/10578 (0.0%)	0.79	8/14394 (0.1%)

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	A	0	1
1	B	0	1
2	I	0	2
2	J	0	2
2	K	0	1
2	L	0	2
All	All	0	9

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	43[A]	SER	CB-OG	-8.82	1.30	1.42
1	A	43[B]	SER	CB-OG	-8.82	1.30	1.42
1	A	43[A]	SER	CA-CB	-6.87	1.42	1.52
1	A	43[B]	SER	CA-CB	-6.87	1.42	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	77	ARG	NE-CZ-NH1	7.62	124.11	120.30
1	A	77	ARG	NE-CZ-NH1	6.74	123.67	120.30
1	A	243	ARG	NE-CZ-NH1	5.79	123.20	120.30
1	B	114	ASP	CB-CG-OD2	-5.76	113.12	118.30
1	B	243	ARG	NE-CZ-NH2	-5.50	117.55	120.30

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	71[A]	GLY	Peptide
1	B	71[B]	GLY	Peptide
2	I	4	PSA	Mainchain,Peptide
2	J	4	PSA	Mainchain,Peptide
2	K	4	PSA	Peptide

## 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	A	2549	0	2449	0	0
1	B	2592	0	2508	6	0
1	C	2539	0	2444	3	0
1	D	2566	0	2464	5	0
2	I	55	0	56	2	0
2	J	55	0	56	0	0
2	K	55	0	56	0	0
2	L	55	0	56	1	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	3	0	0	0	0
4	A	27	0	37	2	0
4	B	20	0	28	3	0
4	C	29	0	39	3	0
4	D	30	0	42	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	15	0	18	3	0
5	D	13	0	17	0	0
6	A	24	0	26	0	0
6	B	12	0	13	0	0
6	C	24	0	26	0	0
6	D	12	0	13	0	0
7	B	16	0	21	0	0
8	A	192	0	0	1	0
8	B	234	0	0	2	2
8	C	134	0	0	1	0
8	D	190	0	0	2	1
8	I	1	0	0	0	0
8	J	2	0	0	0	0
8	K	2	0	0	0	0
8	L	2	0	0	0	0
All	All	11451	0	10369	19	2

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 1.

The worst 5 of 19 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)
1:D:70[A]:LEU:O	8:D:503:HOH:O	1.76	1.01
1:C:5:LEU:HD12	4:C:402:PGE:H22	1.58	0.85
1:B:46[B]:GLN:HG2	8:B:504:HOH:O	1.78	0.82
1:B:5:LEU:HD12	4:B:403:PGE:H52	1.62	0.79
1:B:5:LEU:HD12	4:B:403:PGE:C5	2.25	0.67

All (2) 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)
8:B:720:HOH:O	8:B:720:HOH:O[4_565]	0.97	1.23
8:B:655:HOH:O	8:D:600:HOH:O[11_556]	1.85	0.35

## 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	A	346/339 (102%)	339 (98%)	7 (2%)	0	100	100
1	B	352/339 (104%)	344 (98%)	8 (2%)	0	100	100
1	C	341/339 (101%)	334 (98%)	7 (2%)	0	100	100
1	D	349/339 (103%)	342 (98%)	7 (2%)	0	100	100
2	I	3/7 (43%)	2 (67%)	1 (33%)	0	100	100
2	J	3/7 (43%)	2 (67%)	1 (33%)	0	100	100
2	K	3/7 (43%)	3 (100%)	0	0	100	100
2	L	3/7 (43%)	2 (67%)	1 (33%)	0	100	100
All	All	1400/1384 (101%)	1368 (98%)	32 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	A	286/281 (102%)	285 (100%)	1 (0%)	92	89
1	B	292/281 (104%)	292 (100%)	0	100	100
1	C	284/281 (101%)	283 (100%)	1 (0%)	91	87
1	D	287/281 (102%)	285 (99%)	2 (1%)	84	77
2	I	2/2 (100%)	2 (100%)	0	100	100
2	J	2/2 (100%)	2 (100%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	K	2/2 (100%)	2 (100%)	0	100	100
2	L	2/2 (100%)	2 (100%)	0	100	100
All	All	1157/1132 (102%)	1153 (100%)	4 (0%)	91	89

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	132	SER
1	C	132	SER
1	D	131	THR
1	D	275	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	269	ASN
1	D	272	ASN
1	B	205	ASN
1	C	115	GLN
1	D	46	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](#)

8 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
2	PSA	J	4	2	14,14,15	0.91	0	15,17,19	1.70	6 (40%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	PSA	K	4	2	14,14,15	1.61	3 (21%)	15,17,19	1.44	3 (20%)
2	PSA	L	6	2	14,14,15	1.29	1 (7%)	15,17,19	1.67	3 (20%)
2	PSA	I	6	2	14,14,15	1.55	1 (7%)	15,17,19	1.90	5 (33%)
2	PSA	L	4	2	14,14,15	1.05	0	15,17,19	1.49	4 (26%)
2	PSA	I	4	2	14,14,15	1.04	1 (7%)	15,17,19	1.45	3 (20%)
2	PSA	J	6	2	14,14,15	1.79	4 (28%)	15,17,19	2.72	5 (33%)
2	PSA	K	6	2	14,14,15	1.41	2 (14%)	15,17,19	1.82	2 (13%)

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
2	PSA	J	4	2	-	4/11/11/12	0/1/1/1
2	PSA	K	4	2	-	4/11/11/12	0/1/1/1
2	PSA	L	6	2	-	1/11/11/12	0/1/1/1
2	PSA	I	6	2	-	1/11/11/12	0/1/1/1
2	PSA	L	4	2	-	4/11/11/12	0/1/1/1
2	PSA	I	4	2	-	4/11/11/12	0/1/1/1
2	PSA	J	6	2	-	3/11/11/12	0/1/1/1
2	PSA	K	6	2	-	1/11/11/12	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	6	PSA	O-C	4.40	1.44	1.19
2	J	6	PSA	O-C	4.29	1.44	1.19
2	K	6	PSA	O-C	4.19	1.43	1.19
2	K	4	PSA	CH-CA	-4.14	1.49	1.53
2	L	6	PSA	O-C	3.91	1.42	1.19

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	6	PSA	O-C-CM	-8.42	100.90	125.43
2	K	6	PSA	O-C-CM	-5.12	110.52	125.43
2	I	6	PSA	O-C-CM	-4.60	112.02	125.43
2	L	6	PSA	O-C-CM	-4.26	113.02	125.43

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	6	PSA	CB-CA-CH	3.65	117.79	111.84

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	I	4	PSA	O-C-CM-CH
2	J	4	PSA	O-C-CM-CH
2	L	4	PSA	O-C-CM-CH
2	J	6	PSA	CB-CA-CH-OH
2	K	6	PSA	O-C-CM-CH

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	I	6	PSA	1	0
2	L	4	PSA	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 26 ligands modelled in this entry, 6 are monoatomic - leaving 20 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	PGE	D	404	3	9,9,9	0.46	0	8,8,8	0.34	0
4	PGE	C	403	-	9,9,9	0.46	0	8,8,8	0.66	0
4	PGE	A	404	-	9,9,9	0.37	0	8,8,8	0.54	0
5	1PE	D	407	-	12,12,15	0.40	0	11,11,14	0.60	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	PGE	B	403	-	9,9,9	0.38	0	8,8,8	1.06	0
4	PGE	D	406	3	9,9,9	0.50	0	8,8,8	0.10	0
6	MES	A	406	-	12,12,12	2.16	1 (8%)	14,16,16	1.24	2 (14%)
6	MES	D	408	-	12,12,12	2.29	1 (8%)	14,16,16	2.00	5 (35%)
6	MES	C	406	-	12,12,12	2.40	2 (16%)	14,16,16	1.95	4 (28%)
7	PE3	B	401	-	15,15,42	0.42	0	14,14,41	0.74	0
4	PGE	A	403	-	6,6,9	0.57	0	5,5,8	0.31	0
5	1PE	A	405	-	13,13,15	0.56	0	11,11,14	0.75	0
4	PGE	A	402	3	9,9,9	0.46	0	8,8,8	0.30	0
4	PGE	B	404	3	9,9,9	0.51	0	8,8,8	0.44	0
4	PGE	D	405	-	9,9,9	0.66	0	8,8,8	0.46	0
4	PGE	C	404	3	9,9,9	0.41	0	8,8,8	0.26	0
6	MES	B	405	-	12,12,12	2.08	2 (16%)	14,16,16	1.27	1 (7%)
6	MES	A	407	-	12,12,12	2.07	2 (16%)	14,16,16	1.42	2 (14%)
4	PGE	C	402	-	8,8,9	0.46	0	7,7,8	1.24	1 (14%)
6	MES	C	405	-	12,12,12	2.13	2 (16%)	14,16,16	1.66	3 (21%)

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	PGE	D	404	3	-	2/7/7/7	-
4	PGE	C	403	-	-	5/7/7/7	-
4	PGE	A	404	-	-	3/7/7/7	-
5	1PE	D	407	-	-	1/10/10/13	-
4	PGE	B	403	-	-	0/7/7/7	-
4	PGE	D	406	3	-	1/7/7/7	-
6	MES	A	406	-	-	0/6/14/14	0/1/1/1
6	MES	D	408	-	-	0/6/14/14	0/1/1/1
6	MES	C	406	-	-	1/6/14/14	0/1/1/1
7	PE3	B	401	-	-	4/13/13/40	-
4	PGE	A	403	-	-	3/4/4/7	-
5	1PE	A	405	-	-	2/9/9/13	-
4	PGE	A	402	3	-	2/7/7/7	-
4	PGE	B	404	3	-	1/7/7/7	-
4	PGE	D	405	-	-	3/7/7/7	-
4	PGE	C	404	3	-	3/7/7/7	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	MES	B	405	-	-	0/6/14/14	0/1/1/1
6	MES	A	407	-	-	3/6/14/14	0/1/1/1
4	PGE	C	402	-	-	0/6/6/7	-
6	MES	C	405	-	-	0/6/14/14	0/1/1/1

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	408	MES	C8-S	-7.65	1.66	1.77
6	C	406	MES	C8-S	-7.51	1.66	1.77
6	A	406	MES	C8-S	-7.09	1.67	1.77
6	C	405	MES	C8-S	-6.76	1.67	1.77
6	B	405	MES	C8-S	-6.48	1.68	1.77

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	C	406	MES	O2S-S-C8	4.05	111.79	106.92
6	D	408	MES	O3S-S-C8	3.99	112.21	105.77
6	D	408	MES	O1S-S-C8	3.93	111.65	106.92
6	C	405	MES	O2S-S-C8	3.82	111.51	106.92
6	B	405	MES	O1S-S-C8	3.48	111.10	106.92

There are no chirality outliers.

5 of 34 torsion outliers are listed below:

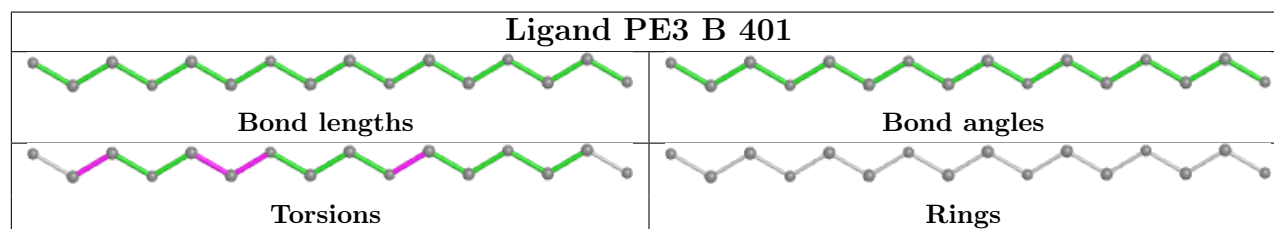
Mol	Chain	Res	Type	Atoms
6	A	407	MES	C7-C8-S-O2S
4	C	403	PGE	O2-C3-C4-O3
4	A	403	PGE	O3-C5-C6-O4
4	A	403	PGE	O2-C3-C4-O3
4	A	404	PGE	O3-C5-C6-O4

There are no ring outliers.

4 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	404	PGE	2	0
4	B	403	PGE	3	0
5	A	405	1PE	3	0
4	C	402	PGE	3	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	339/339 (100%)	0.12	12 (3%) 44 49	25, 32, 47, 68	1 (0%)
1	B	339/339 (100%)	-0.00	10 (2%) 51 56	24, 29, 43, 59	1 (0%)
1	C	339/339 (100%)	0.57	33 (9%) 7 8	30, 40, 65, 110	3 (0%)
1	D	339/339 (100%)	0.18	20 (5%) 22 24	24, 30, 53, 81	1 (0%)
2	I	4/7 (57%)	2.51	1 (25%) 0 0	24, 24, 25, 47	0
2	J	4/7 (57%)	0.75	0 100 100	21, 22, 22, 39	0
2	K	4/7 (57%)	1.69	1 (25%) 0 0	31, 32, 33, 61	0
2	L	4/7 (57%)	0.36	0 100 100	25, 25, 26, 42	0
All	All	1372/1384 (99%)	0.23	77 (5%) 24 27	21, 32, 56, 110	6 (0%)

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	252	TYR	10.0
2	I	7	ALA	9.3
1	C	250	ASP	9.2
1	C	51	VAL	8.9
1	C	249	ARG	7.9

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

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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	PSA	K	6	14/15	0.81	0.14	33,37,54,54	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	PSA	I	6	14/15	0.85	0.12	25,28,41,41	0
2	PSA	J	6	14/15	0.86	0.14	24,25,43,44	0
2	PSA	K	4	14/15	0.90	0.12	28,30,33,34	0
2	PSA	L	6	14/15	0.90	0.11	26,28,43,47	0
2	PSA	J	4	14/15	0.94	0.10	19,20,22,24	0
2	PSA	I	4	14/15	0.95	0.11	22,23,24,26	0
2	PSA	L	4	14/15	0.95	0.09	21,23,26,27	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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.

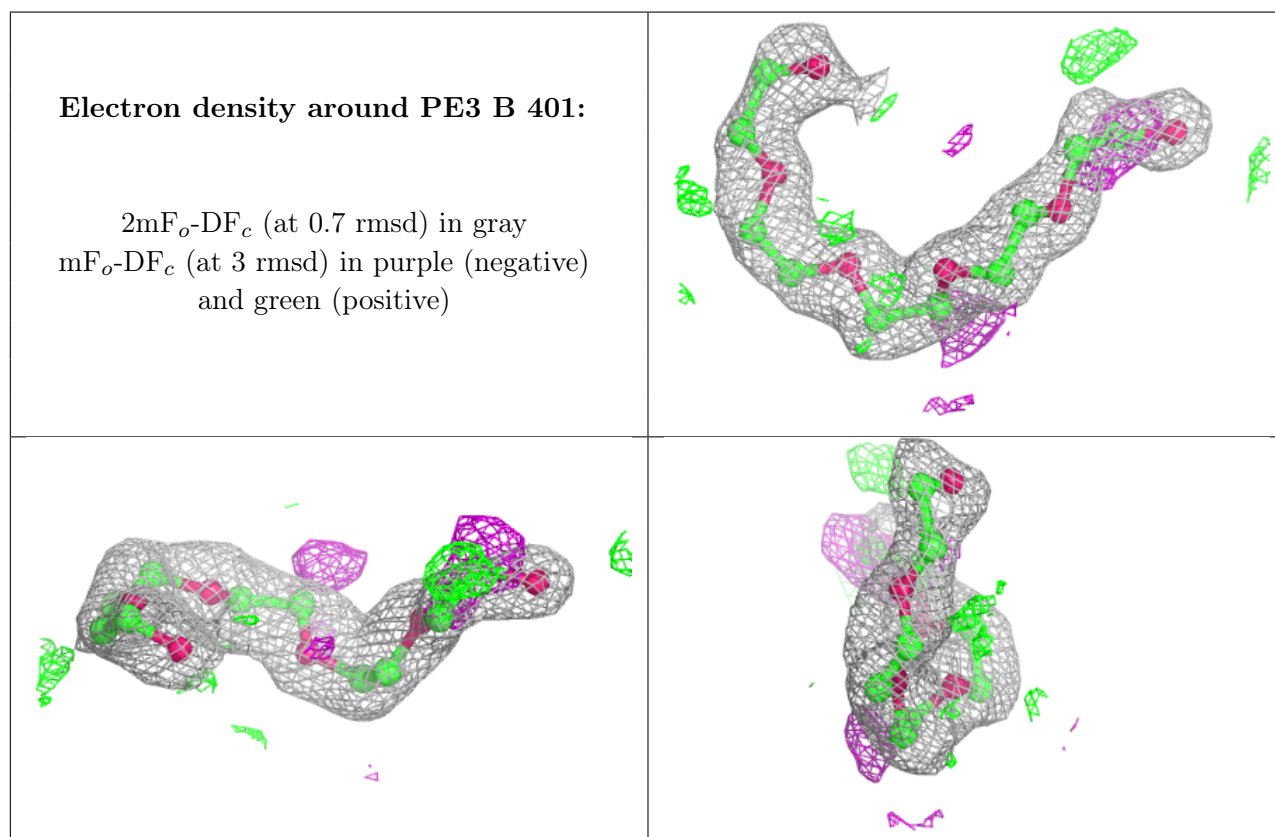
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	PGE	A	403	7/10	0.68	0.18	58,61,64,65	0
3	NA	C	401	1/1	0.69	0.11	72,72,72,72	0
6	MES	C	406	12/12	0.70	0.38	37,56,62,64	1
4	PGE	C	403	10/10	0.73	0.20	50,56,65,66	0
4	PGE	D	404	10/10	0.75	0.20	51,58,62,63	0
4	PGE	D	405	10/10	0.77	0.14	42,47,57,63	0
6	MES	C	405	12/12	0.79	0.37	50,62,69,70	1
3	NA	D	401	1/1	0.80	0.29	60,60,60,60	0
4	PGE	C	404	10/10	0.81	0.15	53,55,60,61	0
3	NA	D	402	1/1	0.82	0.45	59,59,59,59	0
6	MES	A	407	12/12	0.82	0.28	48,61,66,67	1
6	MES	A	406	12/12	0.83	0.24	33,43,50,52	12
4	PGE	D	406	10/10	0.83	0.19	52,56,57,58	0
5	IPE	A	405	15/16	0.84	0.15	29,36,49,57	0
6	MES	D	408	12/12	0.84	0.20	22,32,37,38	12
6	MES	B	405	12/12	0.85	0.30	51,62,64,64	1
4	PGE	A	402	10/10	0.85	0.13	43,45,49,54	0
3	NA	D	403	1/1	0.86	0.26	56,56,56,56	0
4	PGE	B	404	10/10	0.86	0.13	45,47,52,56	0
4	PGE	C	402	9/10	0.87	0.15	35,38,41,43	0
4	PGE	B	403	10/10	0.88	0.17	33,40,45,54	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	PGE	A	404	10/10	0.90	0.12	33,41,53,60	0
7	PE3	B	401	16/43	0.90	0.15	33,40,59,61	0
5	1PE	D	407	13/16	0.95	0.10	28,32,38,44	0
3	NA	B	402	1/1	0.96	0.21	44,44,44,44	0
3	NA	A	401	1/1	0.99	0.09	42,42,42,42	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.