



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 20, 2024 – 12:25 pm GMT

PDB ID : 7AGD  
Title : Protease Sapp1p from *Candida parapsilosis* in complex with KB75  
Authors : Dostal, J.; Heidingsfeld, O.; Brynda, J.  
Deposited on : 2020-09-22  
Resolution : 1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
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

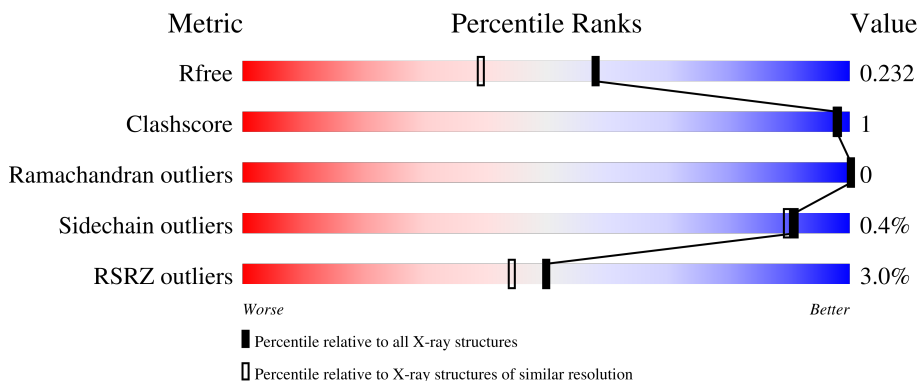
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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))
$R_{free}$	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 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	<div style="display: flex; align-items: center;"> <div style="width: 2%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 98%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 2%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 2%; height: 10px; background-color: orange; margin-right: 2px;"></div> <div style="width: 2%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">2% 98% .</p>
1	B	339	<div style="display: flex; align-items: center;"> <div style="width: 0%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 99%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 1%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 0%; height: 10px; background-color: orange; margin-right: 2px;"></div> <div style="width: 0%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">% 99% .</p>
1	C	339	<div style="display: flex; align-items: center;"> <div style="width: 6%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 93%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 1%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 0%; height: 10px; background-color: orange; margin-right: 2px;"></div> <div style="width: 0%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">6% 99% .</p>
1	D	339	<div style="display: flex; align-items: center;"> <div style="width: 3%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 96%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 1%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 0%; height: 10px; background-color: orange; margin-right: 2px;"></div> <div style="width: 0%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">3% 99% .</p>
2	I	7	<div style="display: flex; align-items: center;"> <div style="width: 43%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 29%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 14%; height: 10px; background-color: orange; margin-right: 2px;"></div> <div style="width: 14%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">43% 29% 14% 14%</p>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	J	7	
2	K	7	
2	L	7	

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 11271 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 2548	C 1589	N 412	O 543	S 4	1	10	0
1	B	339	Total 2588	C 1615	N 420	O 549	S 4	0	15	0
1	C	339	Total 2527	C 1573	N 413	O 537	S 4	0	3	0
1	D	339	Total 2535	C 1580	N 407	O 544	S 4	0	10	0

- Molecule 2 is a protein called KB75.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	I	6	Total 54	C 40	N 5	O 9	0	0	0
2	J	6	Total 54	C 40	N 5	O 9	0	0	0
2	K	6	Total 50	C 38	N 5	O 7	0	0	0
2	L	6	Total 54	C 40	N 5	O 9	0	0	0

- Molecule 3 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C<sub>10</sub>H<sub>22</sub>O<sub>6</sub>).



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

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



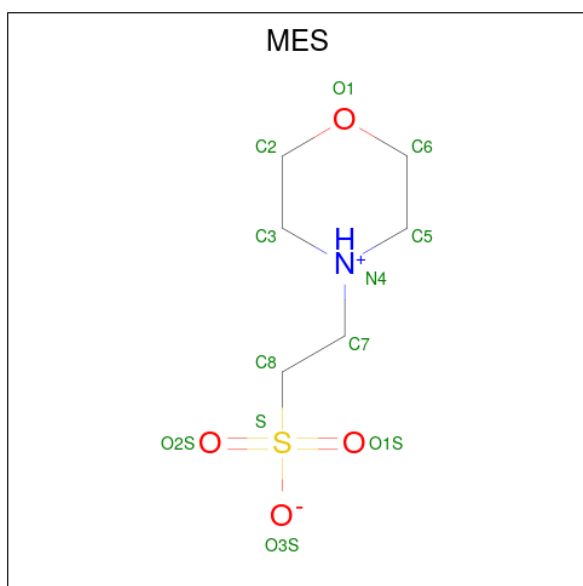
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	C O	0	0
			7	4 3		

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



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

- Molecule 6 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula:  $C_6H_{13}NO_4S$ ).

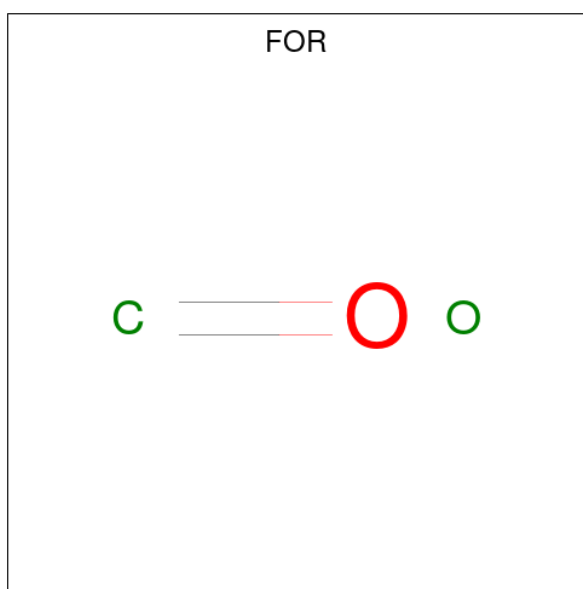


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
6	B	1	12	6	1	4	1	0	0

- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
7	D	1	1	1	0	0

- Molecule 8 is FORMYL GROUP (three-letter code: FOR) (formula: CH<sub>2</sub>O).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
8	J	1	2	1	1	0	0

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
9	A	199	200	200	0	1
9	B	249	251	251	0	2
9	C	129	129	129	0	0
9	D	195	196	196	0	1

*Continued on next page...*

*Continued from previous page...*

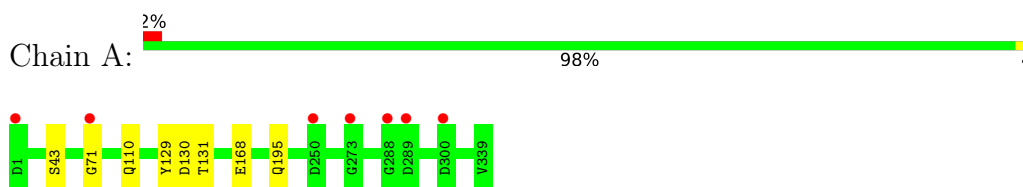
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
9	I	1	Total O 1 1	0	0
9	J	2	Total O 2 2	0	0
9	K	2	Total O 2 2	0	0
9	L	1	Total O 1 1	0	0



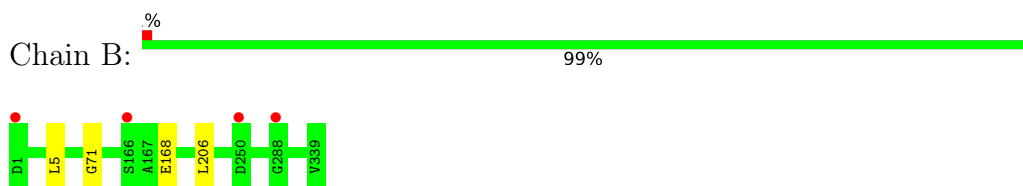
### 3 Residue-property plots [i](#)

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.

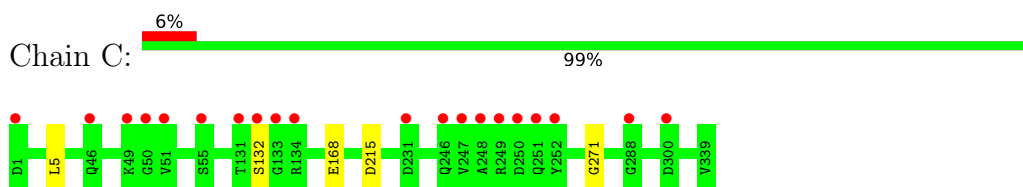
- Molecule 1: Candidapepsin



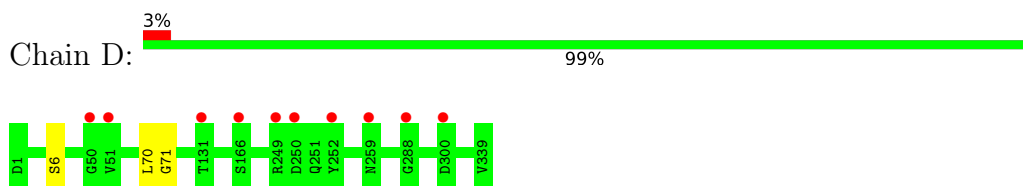
- Molecule 1: Candidapepsin



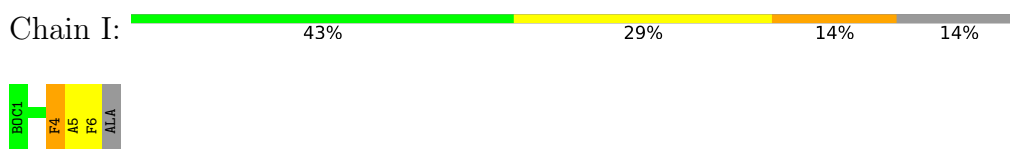
- Molecule 1: Candidapepsin



- Molecule 1: Candidapepsin



- Molecule 2: KB75



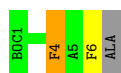
- Molecule 2: KB75

Chain J:  57% 14% 14% 14%



● Molecule 2: KB75

Chain K:  57% 14% 14% 14%



● Molecule 2: KB75

Chain L:  57% 14% 14% 14%



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	172.62Å 172.62Å 253.25Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	18.42 – 1.80 18.42 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.2 (18.42-1.80) 99.3 (18.42-1.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.93 (at 1.80Å)	Xtrriage
Refinement program	REFMAC 5.8.0103	Depositor
R, $R_{free}$	0.201 , 0.223 0.210 , 0.232	Depositor DCC
$R_{free}$ test set	10179 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.9	Xtrriage
Anisotropy	0.005	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 46.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	11271	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	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 41.76 % 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 2.2525e-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: NA, BOC, 1PE, MES, PEG, FOR, PSA, PGE

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	0.68	4/2620 (0.2%)	0.75	3/3568 (0.1%)
1	B	0.60	0/2661	0.73	0/3623
1	C	0.56	0/2579	0.68	1/3513 (0.0%)
1	D	0.60	0/2596	0.72	0/3541
2	I	1.15	0/17	1.05	0/21
2	J	1.01	0/17	0.80	0/21
2	K	1.08	0/17	1.04	0/21
2	L	1.05	0/17	0.77	0/21
All	All	0.61	4/10524 (0.0%)	0.72	4/14329 (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	A	0	1
1	B	0	1
2	I	0	2
2	J	0	2
2	K	0	2
2	L	0	2
All	All	0	10

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	-8.61	1.31	1.42
1	A	43[B]	SER	CB-OG	-8.61	1.31	1.42
1	A	43[A]	SER	CA-CB	7.54	1.64	1.52
1	A	43[B]	SER	CA-CB	7.54	1.64	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	43[A]	SER	CA-CB-OG	-6.51	93.62	111.20
1	A	43[B]	SER	CA-CB-OG	-6.51	93.62	111.20
1	C	271	GLY	N-CA-C	5.96	128.01	113.10
1	A	130	ASP	CB-CG-OD1	5.18	122.96	118.30

There are no chirality outliers.

All (10) 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	Mainchain,Peptide
2	L	4	PSA	Mainchain,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	2548	0	2444	2	0
1	B	2588	0	2500	4	0
1	C	2527	0	2425	2	0
1	D	2535	0	2407	4	0
2	I	54	0	55	1	0
2	J	54	0	55	0	0
2	K	50	0	52	0	0
2	L	54	0	56	0	0
3	A	15	0	18	2	0
3	D	13	0	17	0	0
4	A	7	0	10	1	0
5	B	10	0	14	1	0
5	C	9	0	11	1	0
5	D	10	0	14	0	0
6	B	12	0	13	2	0
7	D	1	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	J	2	0	0	0	0
9	A	200	0	0	1	1
9	B	251	0	0	2	1
9	C	129	0	0	1	0
9	D	196	0	0	2	0
9	I	1	0	0	1	0
9	J	2	0	0	0	0
9	K	2	0	0	0	0
9	L	1	0	0	0	0
All	All	11271	0	10091	15	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.

All (15) 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	9:D:503:HOH:O	1.63	1.13
1:B:5:LEU:HD12	5:B:401:PGE:H52	1.55	0.88
1:C:5:LEU:HD12	5:C:401:PGE:H22	1.55	0.88
1:B:168[A]:GLU:OE2	9:B:502:HOH:O	2.14	0.66
1:A:168:GLU:OE2	9:A:501:HOH:O	2.15	0.64
2:I:5:ALA:O	9:I:101:HOH:O	2.19	0.52
1:C:168:GLU:OE2	9:C:501:HOH:O	2.20	0.49
3:A:401:1PE:C22	1:D:6:SER:H	2.28	0.46
1:D:71[B]:GLY:N	9:D:503:HOH:O	2.45	0.46
1:B:206[B]:LEU:HD23	1:B:206[B]:LEU:HA	1.56	0.45
3:A:401:1PE:H221	1:D:6:SER:H	1.82	0.44
1:B:168[A]:GLU:OE1	9:B:503:HOH:O	2.21	0.44
1:A:129:TYR:CE1	4:A:402:PEG:H22	2.56	0.40

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)
9:B:739:HOH:O	9:B:739:HOH:O[4_565]	0.93	1.27
9:A:676:HOH:O	9:A:676:HOH:O[4_565]	1.78	0.42

## 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	347/339 (102%)	340 (98%)	7 (2%)	0	100	100
1	B	352/339 (104%)	346 (98%)	6 (2%)	0	100	100
1	C	340/339 (100%)	332 (98%)	8 (2%)	0	100	100
1	D	347/339 (102%)	339 (98%)	8 (2%)	0	100	100
2	I	3/7 (43%)	3 (100%)	0	0	100	100
2	J	3/7 (43%)	3 (100%)	0	0	100	100
2	K	3/7 (43%)	2 (67%)	1 (33%)	0	100	100
2	L	3/7 (43%)	3 (100%)	0	0	100	100
All	All	1398/1384 (101%)	1368 (98%)	30 (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%)	283 (99%)	3 (1%)	76	71
1	B	291/281 (104%)	291 (100%)	0	100	100
1	C	282/281 (100%)	280 (99%)	2 (1%)	84	81
1	D	281/281 (100%)	281 (100%)	0	100	100
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	1148/1132 (101%)	1143 (100%)	5 (0%)	91	89

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	110	GLN
1	A	131	THR
1	A	195	GLN
1	C	132	SER
1	C	215	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	115	GLN
1	A	195	GLN
1	B	115	GLN
1	B	205	ASN
1	C	115	GLN
1	D	115	GLN
1	D	272	ASN

### 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	K	4	2	14,14,15	1.64	4 (28%)	15,17,19	1.36	1 (6%)
2	PSA	J	6	8,2	14,14,15	1.83	4 (28%)	15,17,19	3.66	5 (33%)
2	PSA	L	6	2	14,14,15	1.61	2 (14%)	15,17,19	2.58	3 (20%)
2	PSA	I	6	2	14,14,15	1.96	4 (28%)	15,17,19	3.33	7 (46%)
2	PSA	J	4	2	14,14,15	1.39	2 (14%)	15,17,19	1.85	3 (20%)
2	PSA	K	6	2	10,10,15	1.02	1 (10%)	10,12,19	0.66	0
2	PSA	I	4	2	14,14,15	1.43	2 (14%)	15,17,19	1.80	4 (26%)
2	PSA	L	4	2	14,14,15	1.05	1 (7%)	15,17,19	1.86	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	K	4	2	-	2/11/11/12	0/1/1/1
2	PSA	J	6	8,2	-	6/11/11/12	0/1/1/1
2	PSA	L	6	2	-	6/11/11/12	0/1/1/1
2	PSA	I	6	2	-	3/11/11/12	0/1/1/1
2	PSA	J	4	2	-	3/11/11/12	0/1/1/1
2	PSA	K	6	2	-	0/4/4/12	0/1/1/1
2	PSA	I	4	2	-	4/11/11/12	0/1/1/1
2	PSA	L	4	2	-	4/11/11/12	0/1/1/1

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L	6	PSA	O-C	4.36	1.44	1.19
2	I	6	PSA	O-C	4.36	1.44	1.19
2	J	6	PSA	O-C	4.25	1.44	1.19
2	K	4	PSA	CH-CA	-3.83	1.49	1.53
2	I	4	PSA	CH-CA	-3.74	1.49	1.53
2	I	6	PSA	CH-CA	3.62	1.56	1.53
2	J	6	PSA	OH-CH	-3.31	1.36	1.43
2	J	4	PSA	CH-CA	-3.03	1.50	1.53
2	K	4	PSA	O-C	2.89	1.36	1.19
2	I	6	PSA	CM-CH	2.78	1.58	1.53
2	I	6	PSA	CB-CG	2.62	1.57	1.51
2	I	4	PSA	O-C	2.55	1.34	1.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	K	4	PSA	CM-C	2.53	1.56	1.49
2	L	6	PSA	CH-CA	2.51	1.55	1.53
2	J	6	PSA	CB-CG	2.24	1.56	1.51
2	J	6	PSA	CH-CA	2.23	1.55	1.53
2	J	4	PSA	CM-C	2.14	1.55	1.49
2	K	4	PSA	CM-CH	2.08	1.56	1.53
2	L	4	PSA	CM-CH	2.06	1.56	1.53
2	K	6	PSA	CB-CG	2.03	1.56	1.51

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	6	PSA	CB-CA-CH	7.67	124.34	111.84
2	J	6	PSA	O-C-CM	-7.28	104.21	125.43
2	J	6	PSA	CM-CH-CA	6.98	123.91	112.94
2	I	6	PSA	CM-CH-CA	6.78	123.59	112.94
2	L	6	PSA	O-C-CM	-6.44	106.65	125.43
2	L	6	PSA	CM-CH-CA	-6.42	102.86	112.94
2	L	4	PSA	O-C-CM	-5.89	108.25	125.43
2	J	4	PSA	O-C-CM	-5.57	109.20	125.43
2	J	6	PSA	OH-CH-CA	-5.45	99.57	109.40
2	I	6	PSA	CB-CA-CH	5.37	120.59	111.84
2	I	6	PSA	O-C-CM	-5.12	110.51	125.43
2	I	6	PSA	OH-CH-CM	5.05	119.91	109.08
2	I	6	PSA	CH-CM-C	4.78	121.48	113.11
2	I	4	PSA	O-C-CM	-4.45	112.45	125.43
2	K	4	PSA	O-C-CM	-4.01	113.74	125.43
2	I	6	PSA	OH-CH-CA	-2.88	104.20	109.40
2	J	4	PSA	CH-CM-C	-2.59	108.57	113.11
2	I	4	PSA	CZ-CE1-CD1	2.56	124.10	120.19
2	J	4	PSA	OH-CH-CM	2.53	114.51	109.08
2	I	4	PSA	CG-CB-CA	2.48	118.71	113.46
2	L	4	PSA	CH-CM-C	-2.44	108.82	113.11
2	I	6	PSA	CG-CB-CA	-2.43	108.31	113.46
2	L	6	PSA	OH-CH-CA	2.39	113.71	109.40
2	I	4	PSA	CH-CM-C	-2.18	109.28	113.11
2	J	6	PSA	OH-CH-CM	2.02	113.40	109.08

There are no chirality outliers.

All (28) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	I	6	PSA	CA-CH-CM-C
2	I	6	PSA	OH-CH-CM-C
2	J	6	PSA	N-CA-CH-OH
2	J	6	PSA	N-CA-CH-CM
2	J	6	PSA	CB-CA-CH-OH
2	J	6	PSA	CB-CA-CH-CM
2	L	6	PSA	N-CA-CH-OH
2	L	6	PSA	N-CA-CH-CM
2	L	6	PSA	CB-CA-CH-OH
2	L	6	PSA	CB-CA-CH-CM
2	L	6	PSA	O-C-CM-CH
2	L	4	PSA	CA-CH-CM-C
2	J	6	PSA	CA-CH-CM-C
2	J	4	PSA	CA-CB-CG-CD2
2	J	4	PSA	CA-CB-CG-CD1
2	K	4	PSA	CA-CB-CG-CD1
2	K	4	PSA	CA-CB-CG-CD2
2	L	6	PSA	N-CA-CB-CG
2	L	4	PSA	CA-CB-CG-CD2
2	I	4	PSA	CA-CB-CG-CD1
2	L	4	PSA	CA-CB-CG-CD1
2	I	4	PSA	CA-CH-CM-C
2	I	4	PSA	OH-CH-CM-C
2	J	4	PSA	OH-CH-CM-C
2	L	4	PSA	OH-CH-CM-C
2	I	6	PSA	N-CA-CH-OH
2	J	6	PSA	OH-CH-CM-C
2	I	4	PSA	CA-CB-CG-CD2

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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
6	MES	B	402	-	12,12,12	2.26	2 (16%)	14,16,16	2.21	5 (35%)
3	1PE	D	403	-	12,12,15	0.62	0	11,11,14	0.52	0
5	PGE	B	401	-	9,9,9	0.50	0	8,8,8	0.62	0
8	FOR	J	101	2	0,1,1	-	-	-	-	-
5	PGE	C	401	-	8,8,9	0.57	0	7,7,8	1.01	0
4	PEG	A	402	-	6,6,6	0.52	0	5,5,5	0.69	0
5	PGE	D	402	7	9,9,9	0.57	0	8,8,8	0.32	0
3	1PE	A	401	-	13,13,15	0.73	0	11,11,14	0.67	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
6	MES	B	402	-	-	1/6/14/14	0/1/1/1
3	1PE	D	403	-	-	1/10/10/13	-
5	PGE	B	401	-	-	1/7/7/7	-
5	PGE	C	401	-	-	0/6/6/7	-
4	PEG	A	402	-	-	4/4/4/4	-
5	PGE	D	402	7	-	0/7/7/7	-
3	1PE	A	401	-	-	3/9/9/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	402	MES	C8-S	-6.95	1.67	1.77
6	B	402	MES	O2S-S	2.01	1.51	1.45

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	402	MES	O2S-S-C8	4.93	112.86	106.92
6	B	402	MES	O3S-S-C8	3.26	111.04	105.77
6	B	402	MES	O3S-S-O1S	-3.14	103.60	111.27
6	B	402	MES	C6-O1-C2	3.07	120.16	109.89

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	402	MES	C6-C5-N4	2.23	113.48	110.10

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	402	MES	N4-C7-C8-S
3	A	401	1PE	OH2-C12-C22-OH3
4	A	402	PEG	O1-C1-C2-O2
3	D	403	1PE	OH4-C13-C23-OH3
4	A	402	PEG	C1-C2-O2-C3
3	A	401	1PE	C23-C13-OH4-C24
4	A	402	PEG	C4-C3-O2-C2
5	B	401	PGE	C3-C4-O3-C5
3	A	401	1PE	OH4-C13-C23-OH3
4	A	402	PEG	O2-C3-C4-O4

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	402	MES	2	0
5	B	401	PGE	1	0
5	C	401	PGE	1	0
4	A	402	PEG	1	0
3	A	401	1PE	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

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%)	-0.26	7 (2%) 63 59	18, 26, 46, 72	1 (0%)
1	B	339/339 (100%)	-0.42	4 (1%) 79 76	16, 23, 41, 65	1 (0%)
1	C	339/339 (100%)	0.12	20 (5%) 22 17	23, 33, 59, 119	3 (0%)
1	D	339/339 (100%)	-0.27	10 (2%) 51 46	17, 25, 51, 88	1 (0%)
2	I	3/7 (42%)	-0.16	0 100 100	22, 22, 23, 27	0
2	J	3/7 (42%)	-0.47	0 100 100	19, 19, 19, 22	0
2	K	3/7 (42%)	-0.08	0 100 100	30, 30, 30, 35	0
2	L	3/7 (42%)	-0.42	0 100 100	22, 22, 23, 23	0
All	All	1368/1384 (98%)	-0.21	41 (2%) 50 44	16, 26, 49, 119	6 (0%)

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	249	ARG	7.8
1	C	250	ASP	6.3
1	D	250	ASP	5.6
1	C	252	TYR	5.4
1	A	288	GLY	5.0
1	C	51	VAL	4.6
1	C	248	ALA	4.5
1	C	50	GLY	4.0
1	C	49	LYS	3.9
1	C	134	ARG	3.9
1	D	131	THR	3.8
1	C	131	THR	3.8
1	A	250	ASP	3.8
1	C	132	SER	3.6
1	C	288	GLY	3.5
1	D	50	GLY	3.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	51	VAL	3.4
1	D	300	ASP	3.4
1	B	288	GLY	3.4
1	C	246	GLN	3.2
1	A	273	GLY	2.9
1	C	300	ASP	2.8
1	D	166	SER	2.8
1	C	231	ASP	2.7
1	C	251	GLN	2.7
1	D	288	GLY	2.7
1	A	1	ASP	2.6
1	C	247	VAL	2.6
1	C	55	SER	2.6
1	C	46	GLN	2.5
1	D	252	TYR	2.4
1	A	71[A]	GLY	2.4
1	A	300	ASP	2.3
1	A	289	ASP	2.3
1	C	133	GLY	2.3
1	C	1	ASP	2.3
1	D	259	ASN	2.3
1	D	249	ARG	2.3
1	B	250	ASP	2.1
1	B	1	ASP	2.0
1	B	166	SER	2.0

## 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	10/15	0.77	0.17	32,39,41,41	0
2	PSA	I	6	14/15	0.87	0.11	26,29,32,40	0
2	PSA	J	6	14/15	0.88	0.12	24,30,40,44	0
2	PSA	L	6	14/15	0.89	0.13	29,33,44,46	0
2	PSA	K	4	14/15	0.92	0.11	25,28,31,32	0
2	PSA	I	4	14/15	0.96	0.08	20,21,22,22	0
2	PSA	L	4	14/15	0.96	0.07	18,21,23,25	0
2	PSA	J	4	14/15	0.96	0.09	15,19,21,22	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(Å <sup>2</sup> )	Q<0.9
4	PEG	A	402	7/7	0.64	0.23	48,50,51,54	1
6	MES	B	402	12/12	0.74	0.32	35,45,56,60	3
5	PGE	C	401	9/10	0.76	0.27	38,49,51,52	0
3	1PE	A	401	15/16	0.79	0.18	34,46,59,59	0
5	PGE	B	401	10/10	0.83	0.18	45,49,52,53	0
7	NA	D	401	1/1	0.85	0.30	47,47,47,47	0
3	1PE	D	403	13/16	0.87	0.14	36,39,44,47	0
8	FOR	J	101	2/2	0.87	0.17	39,39,39,45	0
5	PGE	D	402	10/10	0.88	0.14	40,46,47,48	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.