



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 09:01 pm GMT

PDB ID : 3MSJ  
Title : Structure of bace (beta secretase) in complex with inhibitor  
Authors : Madden, J.; Kramer, J.; Smith, M.A.; Barker, J.; Godemann, R.  
Deposited on : 2010-04-29  
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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

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  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

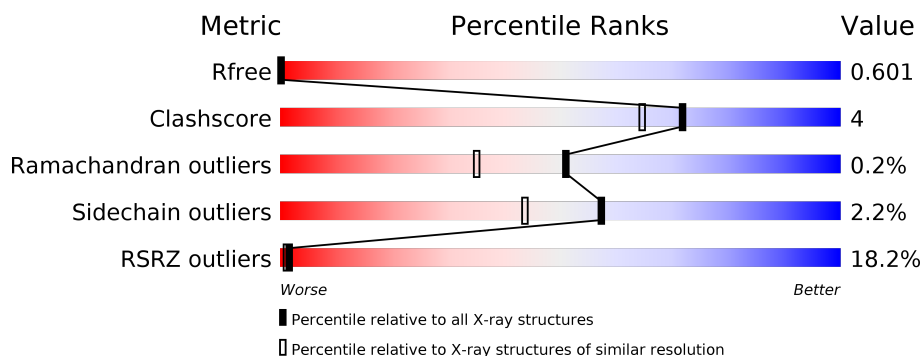
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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))
$R_{free}$	100719	4827 (1.80-1.80)
Clashscore	112137	5742 (1.80-1.80)
Ramachandran outliers	110173	5676 (1.80-1.80)
Sidechain outliers	110143	5675 (1.80-1.80)
RSRZ outliers	101464	4906 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	411	<div> <div>19%</div> <div>78%</div> <div>11%</div> <div>•</div> <div>9%</div> </div>
1	B	411	<div> <div>10%</div> <div>83%</div> <div>7%</div> <div>•</div> <div>9%</div> </div>
1	C	411	<div> <div>20%</div> <div>84%</div> <div>6%</div> <div>•</div> <div>9%</div> </div>

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	GOL	A	395	-	-	-	X
3	GOL	A	396	-	-	-	X
3	GOL	A	399	-	-	-	X
3	GOL	B	394	-	-	X	X
3	GOL	B	395	-	-	-	X
3	GOL	B	397	-	-	-	X
3	GOL	B	398	-	-	-	X
3	GOL	B	400	-	-	-	X
3	GOL	C	394	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9702 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.

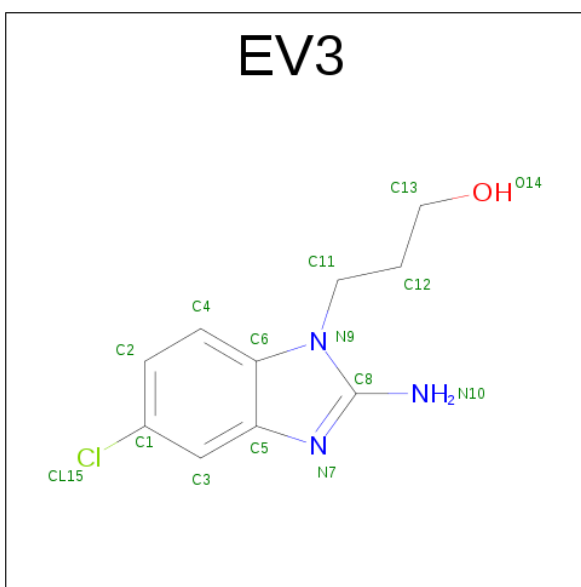
- Molecule 1 is a protein called BETA-SECRETASE 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	373	Total	C	N	O	S	0	4	0
			2962	1897	493	558	14			
1	B	373	Total	C	N	O	S	0	4	0
			2956	1894	493	555	14			
1	C	375	Total	C	N	O	S	0	6	0
			2987	1912	494	567	14			

There are 6 discrepancies between the modelled and reference sequences:

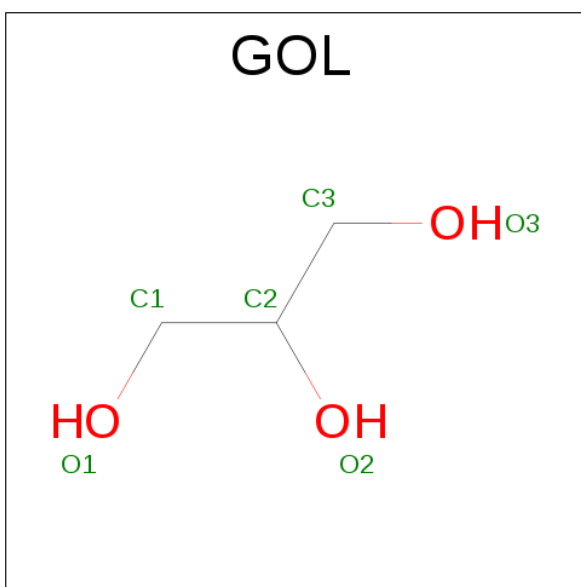
Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	LYS	ARG	ENGINEERED MUTATION	UNP P56817
A	-4	THR	ARG	ENGINEERED MUTATION	UNP P56817
B	-5	LYS	ARG	ENGINEERED MUTATION	UNP P56817
B	-4	THR	ARG	ENGINEERED MUTATION	UNP P56817
C	-5	LYS	ARG	ENGINEERED MUTATION	UNP P56817
C	-4	THR	ARG	ENGINEERED MUTATION	UNP P56817

- Molecule 2 is 3-(2-AMINO-5-CHLORO-1H-BENZIMIDAZOL-1-YL)PROPAN-1-OL (three-letter code: EV3) (formula: C<sub>10</sub>H<sub>12</sub>ClN<sub>3</sub>O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Cl	N	O	0	0
			15	10	1	3	1		
2	B	1	Total	C	Cl	N	O	0	0
			15	10	1	3	1		
2	C	1	Total	C	Cl	N	O	0	0
			15	10	1	3	1		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	236	Total O 236 236	0	0
4	B	243	Total O 246 246	0	3
4	C	167	Total O 168 168	0	1

These plots are drawn for all protein, RNA and DNA 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.

Chain A:

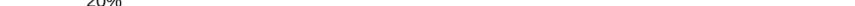
Sequence logo for Chain A. The y-axis represents information content in bits (0.00 to 0.25). The x-axis shows amino acid positions 1 to 40. The logo is color-coded by conservation: red (0.19 bits), green (0.78 bits), yellow (0.11 bits), and grey (0.09 bits). The overall conservation score is 19% (red), 78% (green), 11% (yellow), and 9% (grey).

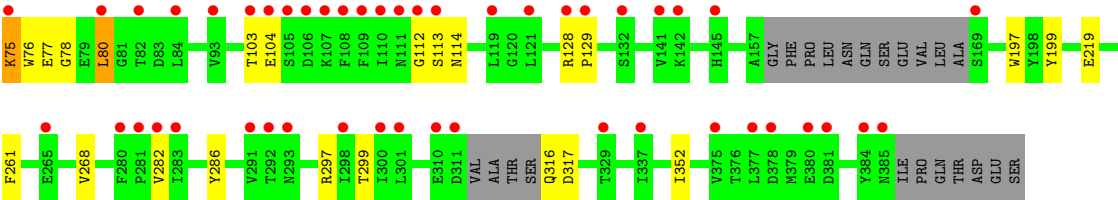
Position	Amino Acid	Information Content (bits)
1	VAL	0.05
2	LEU	0.05
3	ALA	0.05
4	S169	0.05
5	I176	0.05
6	H181	0.05
7	T191	0.05
8	V197	0.05
9	Y198	0.05
10	Y199	0.05
11	E200	0.05
12	V201	0.05
13	I202	0.05
14	I203	0.05
15	V206	0.05
16	Q211	0.05
17	D212	0.05
18	L213	0.05
19	K214	0.05
20	M215	0.05
21	Y222	0.05
22	D223	0.05
23	V227	0.05
24	L234	0.05
25	K238	0.05
26	V239	0.05
27	V240	0.05
28	F241	0.05
29	E242	0.05
30	A243	0.05
31	T248	0.05
32	K256	0.05
33	D259	0.05
34	G260	0.05
35	L267	0.05
36	Q271	0.05
37	A272	0.05
38	V282	0.05
39	I283	0.05
40	S284	0.05
41	L285	0.05
42	V286	0.05
43	L287	0.05
44	N293	0.05
45	R297	0.05
46	I298	0.05
47	T299	0.05
48	I300	0.05
49	V309	0.05
50	E310	0.05
51	ASP	0.05
52	VAL	0.05
53	ALA	0.05
54	THR	0.05
55	SER	0.05
56	Q316	0.05
57	D317	0.05
58	K321	0.05
59	Q326	0.05
60	M333	0.05
61	I337	0.05
62	M338	0.05
63	E339	0.05
64	G340	0.05
65	F341	0.05
66	Y342	0.05
67	V343	0.05
68	V344	0.05
69	F345	0.05
70	I352	0.05
71	G353	0.05
72	F354	0.05
73	S357	0.05
74	H360	0.05
75	E364	0.05
76	T367	0.05
77	T376	0.05
78	L377	0.05
79	D378	0.05
80	N385	0.05
81	I386	0.05
82	PRO	0.05
83	GLN	0.05
84	THR	0.05
85	ASP	0.05
86	GLU	0.05
87	SER	0.05
88	LEU	0.05
89	ASN	0.05
90	GLN	0.05
91	SER	0.05
92	GLN	0.05
93	THR	0.05
94	ASP	0.05
95	LEU	0.05
96	ASN	0.05
97	GLN	0.05
98	SER	0.05
99	GLN	0.05
100	THR	0.05
101	ASP	0.05
102	LEU	0.05
103	ASN	0.05
104	GLN	0.05
105	SER	0.05
106	GLN	0.05
107	THR	0.05
108	ASP	0.05
109	LEU	0.05
110	ASN	0.05
111	GLN	0.05
112	SER	0.05
113	GLN	0.05
114	THR	0.05
115	ASP	0.05
116	LEU	0.05
117	ASN	0.05
118	GLN	0.05
119	SER	0.05
120	GLN	0.05
121	THR	0.05
122	ASP	0.05
123	LEU	0.05
124	ASN	0.05
125	GLN	0.05
126	SER	0.05
127	GLN	0.05
128	THR	0.05
129	ASP	0.05
130	LEU	0.05
131	ASN	0.05
132	GLN	0.05
133	SER	0.05
134	GLN	0.05
135	THR	0.05
136	ASP	0.05
137	LEU	0.05
138	ASN	0.05
139	GLN	0.05
140	SER	

Chain B:

10% 83% 7% 9%

LEU PRO ARG GLU THR ASP GLU PRO GLU PRO GLU PRO GLY LYS T-4 V0 E1 H2 V3 G11 V20 I29 D32 F36 A39 V40 V67 Q73 N92 V95 I99 A100 A101 I102 I110 N111 G112 S113 I118 L119 P129 F136 L140 H145 L152 G156 A157 GLY PHE PRO PRO LEU LEU ASN GLN SER GLU VAL LEU ALA S169 H181 W197 Y198 Y199 I203 V204 R205 Y222 L236 K239 Y256 F261 E265 D266 L267 V268 Q271 A272 F280 D281 V282 I283 S284 L285 N293 I300 Y305 L306 V309 GLU ASP VAL ALA THR SER Q316 D317 I324 T331 V332 M333 I337 V343 T352 G353 F354 T376 L377 D378 M379 E380 Y394 N395 ILE PRO GLN THR ASP GLU SER

Chain C: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	229.20Å 98.83Å 63.13Å 90.00° 103.36° 90.00°	Depositor
Resolution (Å)	111.80 – 1.80 24.03 – 1.80	Depositor EDS
% Data completeness (in resolution range)	94.4 (111.80-1.80) 94.4 (24.03-1.80)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.39 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.177 , 0.208 0.587 , 0.601	Depositor DCC
$R_{free}$ test set	1206 reflections (1.01%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.4	Xtriage
Anisotropy	0.295	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 44.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.024 for -h-2*1,-k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9702	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.45% of the height of the origin peak. No significant pseudotranslation is detected.*

<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: GOL, EV3

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	0.55	0/3046	0.69	2/4136 (0.0%)
1	B	0.55	0/3039	0.71	1/4127 (0.0%)
1	C	0.50	0/3077	0.65	0/4179
All	All	0.53	0/9162	0.68	3/12442 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	67	VAL	CB-CA-C	-6.05	99.90	111.40
1	A	297	ARG	NE-CZ-NH1	5.74	123.17	120.30
1	A	297	ARG	NE-CZ-NH2	-5.10	117.75	120.30

There are no chirality outliers.

There are no planarity outliers.

### 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	2962	0	2875	38	0
1	B	2956	0	2882	21	0
1	C	2987	0	2897	15	0
2	A	15	0	12	3	0
2	B	15	0	12	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	15	0	12	0	0
3	A	42	0	56	2	0
3	B	48	0	64	7	0
3	C	12	0	16	0	0
4	A	236	0	0	3	0
4	B	246	0	0	0	0
4	C	168	0	0	0	0
All	All	9702	0	8826	72	0

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 4.

All (72) 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:A:267:LEU:HD23	1:A:321:LYS:HG3	1.60	0.82
1:C:282:VAL:HG13	1:C:299:THR:HG23	1.68	0.75
1:A:215:MET:HE1	1:A:243:ALA:HB3	1.72	0.69
1:B:0:VAL:H	3:B:394:GOL:H31	1.63	0.63
1:A:309:VAL:HG21	1:A:321:LYS:HD2	1.83	0.61
1:B:0:VAL:HG22	3:B:394:GOL:H32	1.81	0.60
1:C:199:TYR:HB3	1:C:352:ILE:HD11	1.84	0.59
1:A:267:LEU:CD2	1:A:321:LYS:HG3	2.31	0.59
1:C:286:TYR:CZ	1:C:297:ARG:HD3	2.38	0.58
1:A:199:TYR:HB3	1:A:352:ILE:HD11	1.87	0.56
1:C:282:VAL:HG13	1:C:299:THR:CG2	2.34	0.55
1:A:69:VAL:HG21	1:A:76:TRP:CZ2	2.41	0.55
1:B:205:ARG:HA	3:B:400:GOL:H32	1.89	0.55
1:A:2:MET:HE1	1:A:176:ILE:H	1.70	0.55
1:C:67:VAL:HG22	1:C:80:LEU:HD22	1.90	0.54
1:B:199:TYR:HB3	1:B:352:ILE:HD11	1.91	0.52
1:C:69:VAL:HG21	1:C:76:TRP:CZ2	2.45	0.51
1:A:333:MET:HE3	1:A:337:ILE:HG21	1.91	0.51
1:A:215:MET:HE1	1:A:240:VAL:HA	1.93	0.51
1:A:17:GLU:HB3	4:A:607:HOH:O	2.11	0.51
1:A:134:GLU:OE2	1:B:1:GLU:OE2	2.28	0.51
1:A:111:ASN:HD22	1:A:112:GLY:N	2.12	0.47
1:A:213:LEU:HD23	1:A:215:MET:HE3	1.95	0.47
1:A:32:ASP:OD2	1:A:118:ILE:HD11	2.13	0.47
1:B:0:VAL:HG22	3:B:394:GOL:C3	2.45	0.47
2:A:393:EV3:C3	3:A:398:GOL:H2	2.45	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:215:MET:CE	1:A:243:ALA:HB3	2.44	0.47
1:A:357:SER:HB3	1:A:360:HIS:HB3	1.97	0.46
1:C:112:GLY:O	1:C:114:ASN:N	2.48	0.46
1:B:333:MET:HE3	1:B:337:ILE:HG21	1.97	0.46
1:B:95:VAL:HG11	1:B:140:LEU:HD12	1.98	0.46
1:B:156:GLY:O	1:B:157:ALA:CB	2.64	0.46
1:A:364:GLU:HB3	4:A:499:HOH:O	2.16	0.45
1:A:215:MET:CE	1:A:240:VAL:HA	2.47	0.45
1:A:32:ASP:OD2	2:A:393:EV3:N7	2.49	0.45
1:C:77:GLU:HG3	1:C:104:GLU:CG	2.47	0.45
1:A:154:LEU:O	1:A:339:GLU:HA	2.18	0.44
1:B:261:PHE:CD1	1:B:268:VAL:HG23	2.52	0.44
1:C:69:VAL:HG22	1:C:128:ARG:HG3	1.98	0.44
1:A:1:GLU:CD	1:A:1:GLU:H	2.21	0.44
1:A:222:TYR:HA	1:A:223:ASP:HA	1.74	0.44
1:B:11:GLY:N	3:B:397:GOL:H2	2.32	0.44
2:A:393:EV3:C5	3:A:398:GOL:H2	2.48	0.44
1:A:92:ASN:C	1:A:92:ASN:HD22	2.21	0.44
1:B:32:ASP:OD2	2:B:393:EV3:N7	2.51	0.44
1:A:238:LYS:NZ	1:A:242:GLU:OE2	2.49	0.43
1:A:2:MET:CE	1:A:90:GLY:HA2	2.48	0.43
1:B:110:ILE:HB	1:B:113:SER:HB3	2.01	0.43
1:B:293:ASN:HA	1:B:376:THR:O	2.19	0.42
1:A:134:GLU:HB2	3:B:394:GOL:H2	2.01	0.42
1:A:333:MET:CE	1:A:337:ILE:HG21	2.50	0.42
1:B:67:VAL:HG13	1:B:129:PRO:HD3	2.01	0.42
1:A:197:TRP:CG	1:A:198:TYR:N	2.88	0.42
1:C:71:TYR:CE1	1:C:75:LYS:HA	2.55	0.42
1:A:211:GLN:NE2	4:A:483:HOH:O	2.53	0.42
1:A:213:LEU:HG	1:A:215:MET:HE2	2.00	0.42
1:B:92:ASN:HD22	3:B:395:GOL:H31	1.85	0.42
1:A:95:VAL:HG11	1:A:140:LEU:HA	2.02	0.41
1:A:2:MET:CE	1:A:176:ILE:H	2.33	0.41
1:A:238:LYS:HE2	1:A:326:GLN:HE22	1.84	0.41
1:A:82:THR:HG21	1:B:0:VAL:HG12	2.02	0.41
1:A:42:ALA:CB	1:A:101:ALA:HB1	2.50	0.41
1:A:63:LEU:O	1:B:181:HIS:HD2	2.04	0.41
1:C:65:LYS:NZ	1:C:129:PRO:HG3	2.35	0.41
1:B:267:LEU:HD21	1:B:309:VAL:HG21	2.02	0.41
1:B:32:ASP:OD2	1:B:118:ILE:HD11	2.21	0.41
1:C:78:GLY:HA2	1:C:103:THR:HG22	2.03	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:222:TYR:OH	1:C:219:GLU:OE1	2.26	0.41
1:C:63:LEU:HD12	1:C:80:LEU:HB3	2.02	0.41
1:A:293:ASN:HA	1:A:376:THR:O	2.21	0.40
1:C:261:PHE:CD1	1:C:268:VAL:HG23	2.57	0.40

There are no symmetry-related clashes.

## 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	371/411 (90%)	362 (98%)	9 (2%)	0	100	100
1	B	371/411 (90%)	364 (98%)	7 (2%)	0	100	100
1	C	375/411 (91%)	364 (97%)	9 (2%)	2 (0%)	32	17
All	All	1117/1233 (91%)	1090 (98%)	25 (2%)	2 (0%)	51	35

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	-3	GLY
1	C	113	SER

### 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	322/352 (92%)	313 (97%)	9 (3%)	49	34
1	B	322/352 (92%)	317 (98%)	5 (2%)	68	58
1	C	326/352 (93%)	319 (98%)	7 (2%)	59	46
All	All	970/1056 (92%)	949 (98%)	21 (2%)	57	44

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	17	GLU
1	A	72	THR
1	A	92	ASN
1	A	111	ASN
1	A	169	SER
1	A	197	TRP
1	A	259	ASP
1	A	364	GLU
1	A	367	THR
1	B	67	VAL
1	B	73	GLN
1	B	111	ASN
1	B	197	TRP
1	B	239	LYS
1	C	65	LYS
1	C	67	VAL
1	C	75	LYS
1	C	80	LEU
1	C	197	TRP
1	C	316	GLN
1	C	317	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (20) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	28	ASN
1	A	73	GLN
1	A	92	ASN
1	A	111	ASN
1	A	114	ASN
1	A	211	GLN
1	A	266	GLN
1	A	326	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	362	HIS
1	B	28	ASN
1	B	53	GLN
1	B	92	ASN
1	B	111	ASN
1	B	181	HIS
1	B	362	HIS
1	C	28	ASN
1	C	53	GLN
1	C	73	GLN
1	C	92	ASN
1	C	362	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

20 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	$\# Z  > 2$
2	EV3	A	393	-	14,16,16	1.45	2 (14%)	15,22,22	0.99	1 (6%)
3	GOL	A	394	-	5,5,5	0.35	0	5,5,5	0.24	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GOL	A	395	-	5,5,5	0.61	0	5,5,5	0.44	0
3	GOL	A	396	-	5,5,5	0.50	0	5,5,5	0.56	0
3	GOL	A	397	-	5,5,5	0.31	0	5,5,5	0.32	0
3	GOL	A	398	-	5,5,5	0.34	0	5,5,5	0.37	0
3	GOL	A	399	-	5,5,5	0.36	0	5,5,5	0.71	0
3	GOL	A	400	-	5,5,5	0.36	0	5,5,5	0.15	0
2	EV3	B	393	-	14,16,16	1.47	2 (14%)	15,22,22	1.00	0
3	GOL	B	394	-	5,5,5	0.68	0	5,5,5	1.51	2 (40%)
3	GOL	B	395	-	5,5,5	0.31	0	5,5,5	0.30	0
3	GOL	B	396	-	5,5,5	0.40	0	5,5,5	0.54	0
3	GOL	B	397	-	5,5,5	0.26	0	5,5,5	0.40	0
3	GOL	B	398	-	5,5,5	0.35	0	5,5,5	0.36	0
3	GOL	B	399	-	5,5,5	0.34	0	5,5,5	0.50	0
3	GOL	B	400	-	5,5,5	0.32	0	5,5,5	0.55	0
3	GOL	B	401	-	5,5,5	0.32	0	5,5,5	0.49	0
2	EV3	C	393	-	14,16,16	1.48	2 (14%)	15,22,22	0.90	0
3	GOL	C	394	-	5,5,5	0.40	0	5,5,5	0.47	0
3	GOL	C	395	-	5,5,5	0.39	0	5,5,5	0.33	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
2	EV3	A	393	-	-	0/4/4/4	0/2/2/2
3	GOL	A	394	-	-	0/4/4/4	0/0/0/0
3	GOL	A	395	-	-	0/4/4/4	0/0/0/0
3	GOL	A	396	-	-	0/4/4/4	0/0/0/0
3	GOL	A	397	-	-	0/4/4/4	0/0/0/0
3	GOL	A	398	-	-	0/4/4/4	0/0/0/0
3	GOL	A	399	-	-	0/4/4/4	0/0/0/0
3	GOL	A	400	-	-	0/4/4/4	0/0/0/0
2	EV3	B	393	-	-	0/4/4/4	0/2/2/2
3	GOL	B	394	-	-	0/4/4/4	0/0/0/0
3	GOL	B	395	-	-	0/4/4/4	0/0/0/0
3	GOL	B	396	-	-	0/4/4/4	0/0/0/0
3	GOL	B	397	-	-	0/4/4/4	0/0/0/0
3	GOL	B	398	-	-	0/4/4/4	0/0/0/0
3	GOL	B	399	-	-	0/4/4/4	0/0/0/0
3	GOL	B	400	-	-	0/4/4/4	0/0/0/0
3	GOL	B	401	-	-	0/4/4/4	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EV3	C	393	-	-	0/4/4/4	0/2/2/2
3	GOL	C	394	-	-	0/4/4/4	0/0/0/0
3	GOL	C	395	-	-	0/4/4/4	0/0/0/0

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	393	EV3	C3-C5	-2.61	1.37	1.41
2	C	393	EV3	C3-C5	-2.48	1.38	1.41
2	B	393	EV3	C3-C5	-2.15	1.38	1.41
2	A	393	EV3	C8-N10	3.02	1.40	1.33
2	C	393	EV3	C8-N10	3.19	1.40	1.33
2	B	393	EV3	C8-N10	3.30	1.40	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	393	EV3	C1-C3-C5	-2.03	117.73	119.52
3	B	394	GOL	O3-C3-C2	2.08	120.56	110.07
3	B	394	GOL	O1-C1-C2	2.28	121.57	110.07

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	393	EV3	3	0
3	A	398	GOL	2	0
2	B	393	EV3	1	0
3	B	394	GOL	4	0
3	B	395	GOL	1	0
3	B	397	GOL	1	0
3	B	400	GOL	1	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	373/411 (90%)	1.03	80 (21%) 1 1	9, 20, 34, 179	5 (1%)
1	B	373/411 (90%)	0.58	43 (11%) 5 4	6, 24, 74, 100	4 (1%)
1	C	375/411 (91%)	1.23	81 (21%) 1 1	11, 20, 37, 618	4 (1%)
All	All	1121/1233 (90%)	0.95	204 (18%) 1 1	6, 21, 58, 618	13 (1%)

All (204) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	48	LEU	15.2
1	A	48	LEU	9.3
1	C	45	HIS	7.4
1	A	72	THR	7.3
1	C	73	GLN	7.1
1	C	72	THR	6.8
1	A	73	GLN	6.6
1	C	80	LEU	6.4
1	C	46	PRO	6.1
1	C	71	TYR	5.6
1	C	110	ILE	5.4
1	C	70	PRO	5.3
1	C	145	HIS	5.3
1	A	43	ALA	5.3
1	A	47	PHE	5.2
1	C	47	PHE	5.2
1	A	111	ASN	5.1
1	C	64	ARG	4.9
1	C	74	GLY	4.8
1	A	337	ILE	4.8
1	C	63	LEU	4.7
1	A	310	GLU	4.6
1	A	-2	SER	4.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	132	SER	4.5
1	B	272	ALA	4.5
1	C	103	THR	4.4
1	A	298	ILE	4.4
1	A	112	GLY	4.4
1	A	203	ILE	4.3
1	A	285	LEU	4.3
1	C	377	LEU	4.3
1	A	201	VAL	4.3
1	C	55	GLN	4.2
1	A	119	LEU	4.2
1	A	110	ILE	4.1
1	A	49	HIS	4.1
1	A	56	LEU	4.1
1	A	145	HIS	4.0
1	A	352	ILE	4.0
1	C	112	GLY	4.0
1	A	0	VAL	4.0
1	C	75	LYS	4.0
1	C	49	HIS	4.0
1	C	68	TYR	4.0
1	C	291	VAL	4.0
1	C	142	LYS	3.9
1	C	111	ASN	3.9
1	C	113	SER	3.9
1	A	10	SER	3.8
1	A	9	LYS	3.8
1	B	265	GLU	3.8
1	A	316	GLN	3.8
1	A	23	PRO	3.7
1	A	51	TYR	3.7
1	A	55	GLN	3.7
1	A	354	PHE	3.7
1	C	58	SER	3.7
1	B	38	PHE	3.6
1	C	378	ASP	3.6
1	A	92	ASN	3.6
1	C	311	ASP	3.6
1	A	74	GLY	3.6
1	A	68	TYR	3.5
1	C	119	LEU	3.5
1	A	-3	GLY	3.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	287	LEU	3.5
1	A	169	SER	3.4
1	A	343	VAL	3.4
1	A	45	HIS	3.4
1	C	380	GLU	3.4
1	C	84	LEU	3.4
1	C	298	ILE	3.4
1	B	380	GLU	3.4
1	A	106	ASP	3.4
1	C	43	ALA	3.3
1	B	145	HIS	3.3
1	B	73	GLN	3.3
1	B	384	TYR	3.3
1	C	44	PRO	3.3
1	B	39	ALA	3.3
1	B	99	ILE	3.3
1	C	265	GLU	3.2
1	C	56	LEU	3.2
1	C	106	ASP	3.2
1	C	107	LYS	3.2
1	C	310	GLU	3.2
1	A	38	PHE	3.2
1	C	23	PRO	3.2
1	C	24	PRO	3.1
1	A	283	ILE	3.1
1	A	42	ALA	3.1
1	C	300	ILE	3.1
1	A	300	ILE	3.1
1	A	191	THR	3.1
1	A	-1	PHE	3.0
1	C	51	TYR	3.0
1	C	109	PHE	3.0
1	C	31	VAL	3.0
1	C	93	VAL	3.0
1	A	1	GLU	3.0
1	C	104	GLU	2.9
1	C	292	THR	2.9
1	C	60	TYR	2.9
1	C	381	ASP	2.9
1	B	337	ILE	2.9
1	A	71	TYR	2.9
1	C	25	GLN	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	306	LEU	2.9
1	C	66	GLY	2.9
1	A	271	GLN	2.8
1	A	227	VAL	2.8
1	B	40	VAL	2.8
1	A	104	GLU	2.8
1	C	385	ASN	2.8
1	C	141	VAL	2.8
1	A	121	LEU	2.8
1	B	157	ALA	2.8
1	A	202	ILE	2.8
1	B	203	ILE	2.8
1	A	107	LYS	2.8
1	B	378	ASP	2.8
1	B	271	GLN	2.8
1	B	305	TYR	2.8
1	C	384	TYR	2.8
1	B	100	ALA	2.7
1	C	375	VAL	2.7
1	C	21	GLY	2.7
1	B	281	PRO	2.7
1	C	22	SER	2.7
1	B	29	ILE	2.7
1	A	333	MET	2.7
1	C	337	ILE	2.7
1	A	150	PHE	2.7
1	A	44	PRO	2.6
1	C	50	ARG	2.6
1	A	344	VAL	2.6
1	B	343	VAL	2.6
1	B	152	LEU	2.6
1	B	285	LEU	2.6
1	A	118	ILE	2.6
1	A	378	ASP	2.5
1	B	169	SER	2.5
1	B	324	ILE	2.5
1	A	206	VAL	2.5
1	B	0	VAL	2.5
1	C	282	VAL	2.5
1	A	152	LEU	2.5
1	B	331	THR	2.5
1	B	317	ASP	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	102	ILE	2.5
1	B	280	PHE	2.5
1	A	317	ASP	2.5
1	A	234	LEU	2.5
1	A	156	GLY	2.4
1	C	301	LEU	2.4
1	C	39	ALA	2.4
1	A	53	GLN	2.4
1	C	42	ALA	2.4
1	C	121	LEU	2.4
1	B	136	PHE	2.4
1	C	108	PHE	2.4
1	A	99	ILE	2.4
1	B	300	ILE	2.4
1	C	281	PRO	2.4
1	B	101	ALA	2.3
1	B	283	ILE	2.3
1	C	53	GLN	2.3
1	A	4	ASP	2.3
1	B	20[A]	VAL	2.3
1	A	46	PRO	2.3
1	A	248	ILE	2.3
1	C	129	PRO	2.3
1	A	20	VAL	2.3
1	C	283	ILE	2.3
1	C	329	THR	2.3
1	A	256	LYS	2.3
1	B	354	PHE	2.2
1	A	385	ASN	2.2
1	C	293	ASN	2.2
1	B	119	LEU	2.2
1	A	272	ALA	2.2
1	A	282	VAL	2.2
1	A	113	SER	2.2
1	A	345	PHE	2.2
1	B	256	LYS	2.2
1	A	260	GLY	2.1
1	B	118	ILE	2.1
1	A	39	ALA	2.1
1	B	3	VAL	2.1
1	A	12	GLN	2.1
1	B	316	GLN	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	82	THR	2.1
1	A	341	PHE	2.1
1	C	280	PHE	2.1
1	B	375	VAL	2.1
1	B	236	LEU	2.1
1	A	26	THR	2.1
1	C	128	ARG	2.1
1	C	69	VAL	2.1
1	C	169	SER	2.0
1	A	109	PHE	2.0
1	C	65	LYS	2.0
1	C	62	ASP	2.0
1	C	105	SER	2.0

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

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	GOL	A	395	6/6	0.81	0.28	9.88	21,31,34,42	0
3	GOL	B	400	6/6	0.74	0.36	7.40	35,42,43,46	0
3	GOL	A	399	6/6	0.69	0.32	7.11	51,56,57,58	0
3	GOL	B	395	6/6	0.81	0.24	5.40	36,41,44,44	0
3	GOL	B	394	6/6	0.83	0.33	4.54	5,19,31,38	0
3	GOL	B	398	6/6	0.81	0.35	3.74	19,39,42,44	0
3	GOL	C	394	6/6	0.86	0.22	3.18	21,34,35,43	0
3	GOL	B	397	6/6	0.88	0.21	2.37	29,33,33,35	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	GOL	A	396	6/6	0.64	0.23	2.16	22,40,41,45	0
3	GOL	B	396	6/6	0.87	0.21	1.59	11,37,42,42	0
3	GOL	B	401	6/6	0.78	0.22	0.95	51,54,56,58	0
2	EV3	B	393	15/15	0.79	0.17	0.50	38,42,44,44	0
2	EV3	C	393	15/15	0.66	0.21	0.24	62,72,72,72	0
2	EV3	A	393	15/15	0.84	0.13	-0.01	27,31,34,36	0
3	GOL	B	399	6/6	0.78	0.16	-0.20	31,40,42,43	0
3	GOL	A	398	6/6	0.68	0.14	-0.21	50,53,55,58	0
3	GOL	A	400	6/6	0.74	0.21	-	51,52,54,56	0
3	GOL	A	397	6/6	0.84	0.15	-	47,47,48,49	0
3	GOL	A	394	6/6	0.76	0.17	-	46,51,52,53	0
3	GOL	C	395	6/6	0.78	0.26	-	43,47,49,52	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.