Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 10:20 PM GMT

PDB ID	:	1T6B
Title	:	Crystal structure of B. anthracis Protective Antigen complexed with human Anthrax toxin receptor
Authors	:	Santelli, E.; Bankston, L.A.; Leppla, S.H.; Liddington, R.C.
Deposited on	:	2004-05-05
Resolution	:	2.50 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

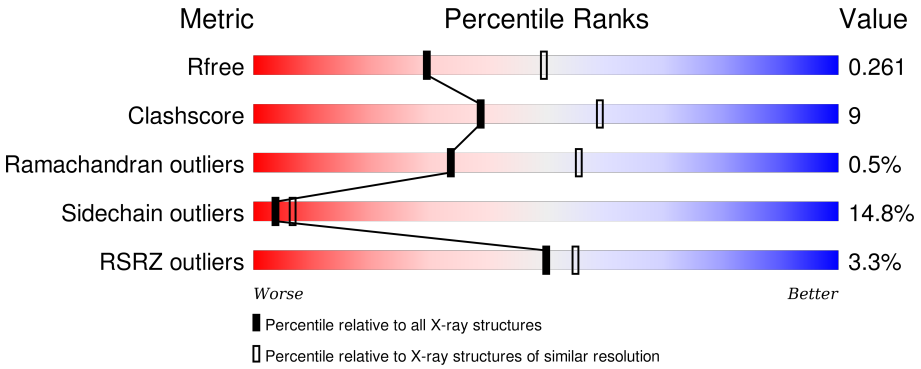
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.50 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}	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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 >=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 <=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	X	735	3% 66% 21% • 8%
2	Y	189	3% 68% 19% • 10%

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
5	NA	X	803	-	-	-	X
6	PG4	X	804	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6852 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 Protective Antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	676	Total	C	N	O	S	0	0	0
			5379	3378	914	1079	8			

- Molecule 2 is a protein called Anthrax toxin receptor 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	Y	170	Total	C	N	O	S	0	0	0
			1316	846	220	248	2			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Y	30	GLY	-	CLONING ARTIFACT	UNP P58335
Y	31	SER	-	CLONING ARTIFACT	UNP P58335
Y	32	HIS	-	CLONING ARTIFACT	UNP P58335
Y	33	MET	-	CLONING ARTIFACT	UNP P58335
Y	34	LEU	-	CLONING ARTIFACT	UNP P58335
Y	36	ASP	GLN	CLONING ARTIFACT	UNP P58335
Y	38	ARG	SER	CLONING ARTIFACT	UNP P58335
Y	39	GLY	CYS	CLONING ARTIFACT	UNP P58335
Y	213	SER	-	CLONING ARTIFACT	UNP P58335
Y	214	ASN	-	CLONING ARTIFACT	UNP P58335
Y	215	GLY	-	CLONING ARTIFACT	UNP P58335
Y	216	ILE	-	CLONING ARTIFACT	UNP P58335
Y	217	ALA	-	CLONING ARTIFACT	UNP P58335
Y	218	ALA	-	CLONING ARTIFACT	UNP P58335

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	X	2	Total	Ca	0	0
			2	2		

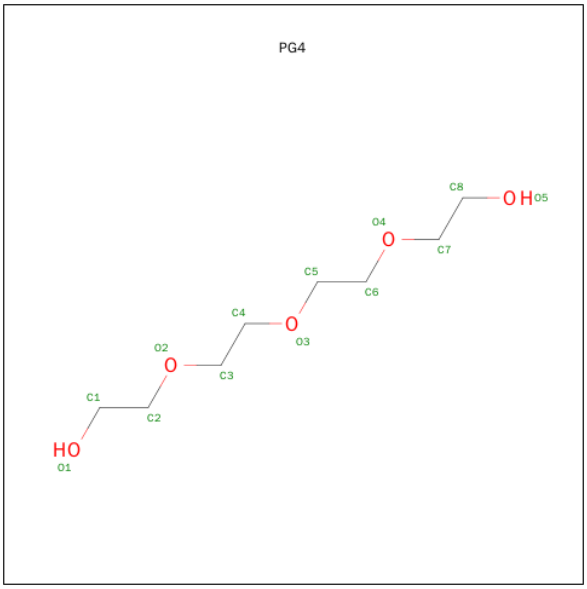
- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	Y	1	Total Mn 1 1	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	X	2	Total Na 2 2	0	0

- Molecule 6 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	X	1	Total C O 13 8 5	0	0

- Molecule 7 is water.

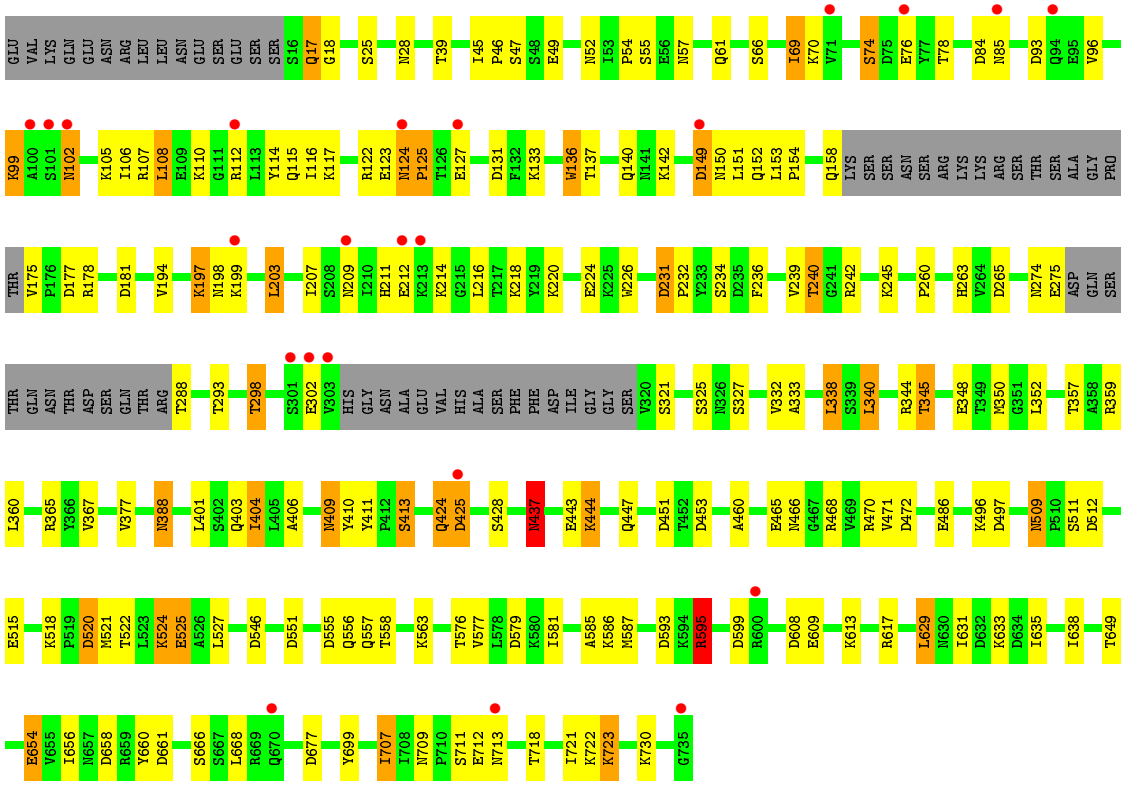
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	X	119	Total O 119 119	0	0
7	Y	20	Total O 20 20	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

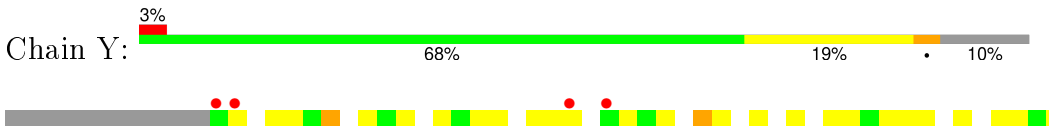
- Molecule 1: Protective Antigen

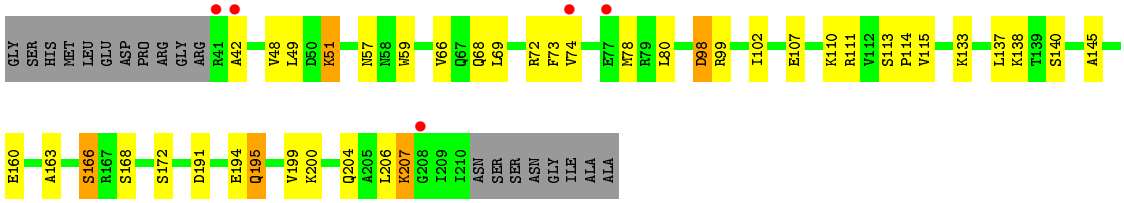
Chain X:



- Molecule 2: Anthrax toxin receptor 2

Chain Y:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	88.16Å 94.12Å 135.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.50 29.56 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.00-2.50) 99.7 (29.56-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.18	Depositor
$< I/\sigma(I) >$ [1]	2.60 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.207 , 0.266 0.202 , 0.261	Depositor DCC
R_{free} test set	1990 reflections (5.29%)	DCC
Wilson B-factor (Å^2)	33.6	Xtriage
Anisotropy	0.441	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 39.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 39610 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6852	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.44% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, CA, MN, PG4

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	X	0.80	2/5474 (0.0%)	0.95	23/7418 (0.3%)
2	Y	0.79	0/1337	0.83	1/1801 (0.1%)
All	All	0.80	2/6811 (0.0%)	0.93	24/9219 (0.3%)

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
2	Y	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	X	437	ASN	CB-CG	5.86	1.64	1.51
1	X	525	GLU	CG-CD	5.73	1.60	1.51

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	X	512	ASP	CB-CG-OD2	8.75	126.17	118.30
1	X	595	ARG	NE-CZ-NH1	7.43	124.02	120.30
1	X	520	ASP	CB-CG-OD2	7.37	124.93	118.30
2	Y	98	ASP	CB-CG-OD2	7.29	124.86	118.30
1	X	231	ASP	CB-CG-OD2	7.18	124.77	118.30
1	X	608	ASP	CB-CG-OD2	6.97	124.57	118.30
1	X	131	ASP	CB-CG-OD2	6.65	124.28	118.30
1	X	425	ASP	CB-CG-OD2	6.53	124.18	118.30
1	X	265	ASP	CB-CG-OD2	6.39	124.06	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	X	555	ASP	CB-CG-OD2	6.36	124.02	118.30
1	X	472	ASP	CB-CG-OD2	6.35	124.01	118.30
1	X	84	ASP	CB-CG-OD2	6.34	124.00	118.30
1	X	658	ASP	CB-CG-OD2	6.28	123.95	118.30
1	X	177	ASP	CB-CG-OD2	6.09	123.78	118.30
1	X	595	ARG	NE-CZ-NH2	-5.89	117.35	120.30
1	X	677	ASP	CB-CG-OD2	5.84	123.56	118.30
1	X	451	ASP	CB-CG-OD2	5.60	123.34	118.30
1	X	365	ARG	NE-CZ-NH2	-5.56	117.52	120.30
1	X	497	ASP	CB-CG-OD2	5.33	123.10	118.30
1	X	203	LEU	CB-CG-CD2	-5.27	102.05	111.00
1	X	551	ASP	CB-CG-OD2	5.26	123.03	118.30
1	X	546	ASP	CB-CG-OD2	5.21	122.98	118.30
1	X	661	ASP	CB-CG-OD2	5.18	122.97	118.30
1	X	181	ASP	CB-CG-OD1	5.03	122.82	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	Y	42	ALA	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	X	5379	0	5313	106	0
2	Y	1316	0	1348	21	0
3	X	2	0	0	0	0
4	Y	1	0	0	0	0
5	X	2	0	0	0	0
6	X	13	0	18	1	0
7	X	119	0	0	20	0
7	Y	20	0	0	2	0
All	All	6852	0	6679	120	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 9.

All (120) 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:X:409:ASN:HD22	1:X:410:TYR:H	1.12	0.97
1:X:327:SER:HA	7:X:915:HOH:O	1.61	0.96
1:X:521:MET:HE1	1:X:525:GLU:HB3	1.45	0.94
1:X:521:MET:CE	1:X:525:GLU:HB3	1.99	0.93
1:X:425:ASP:HA	7:X:869:HOH:O	1.69	0.90
1:X:123:GLU:O	1:X:125:PRO:HD3	1.75	0.84
1:X:345:THR:HG22	1:X:443:GLU:OE2	1.79	0.82
1:X:496:LYS:HB3	7:X:913:HOH:O	1.82	0.80
1:X:654:GLU:OE1	2:Y:51:LYS:NZ	2.21	0.73
1:X:236:PHE:O	1:X:240:THR:HB	1.90	0.72
1:X:629:LEU:HD12	1:X:631:ILE:HD11	1.71	0.72
1:X:656:ILE:HD12	2:Y:115:VAL:HG11	1.72	0.72
1:X:175:VAL:HA	7:X:839:HOH:O	1.93	0.68
1:X:338:LEU:O	1:X:338:LEU:HD23	1.93	0.68
1:X:240:THR:HG21	1:X:242:ARG:HH11	1.59	0.68
1:X:152:GLN:HB3	7:X:827:HOH:O	1.95	0.67
1:X:521:MET:HE2	1:X:522:THR:O	1.95	0.66
1:X:298:THR:HG22	7:X:902:HOH:O	1.96	0.66
1:X:263:HIS:HD2	7:X:877:HOH:O	1.79	0.63
1:X:46:PRO:HD2	1:X:49:GLU:OE2	2.01	0.61
1:X:593:ASP:OD1	1:X:595:ARG:HG2	2.02	0.60
1:X:231:ASP:HB2	1:X:232:PRO:HD3	1.82	0.60
1:X:437:ASN:HB3	7:X:891:HOH:O	2.01	0.60
1:X:327:SER:CA	7:X:915:HOH:O	2.35	0.60
1:X:226:TRP:CE2	1:X:234:SER:HB3	2.36	0.60
1:X:509:ASN:HD22	1:X:511:SER:H	1.50	0.58
1:X:194:VAL:HG11	1:X:239:VAL:HB	1.85	0.58
1:X:218:LYS:HE2	1:X:220:LYS:HE3	1.84	0.58
1:X:18:GLY:O	1:X:153:LEU:HD23	2.04	0.58
1:X:404:ILE:HD11	1:X:406:ALA:HB2	1.85	0.58
1:X:635:ILE:O	1:X:638:ILE:HG13	2.04	0.57
1:X:558:THR:HG23	1:X:587:MET:HG2	1.85	0.57
1:X:409:ASN:HD22	1:X:410:TYR:N	1.93	0.56
1:X:524:LYS:HE2	1:X:579:ASP:HB3	1.86	0.56
1:X:345:THR:CG2	1:X:443:GLU:OE2	2.53	0.56
1:X:350:MET:CE	1:X:352:LEU:HD13	2.35	0.56
1:X:403:GLN:HE22	1:X:413:SER:H	1.54	0.56
1:X:437:ASN:CB	7:X:891:HOH:O	2.53	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:X:298:THR:CG2	7:X:902:HOH:O	2.52	0.55
2:Y:73:PHE:HB3	2:Y:78:MET:SD	2.46	0.55
1:X:656:ILE:HD12	2:Y:115:VAL:CG1	2.37	0.55
1:X:345:THR:HG22	1:X:443:GLU:CD	2.27	0.54
1:X:656:ILE:HD12	2:Y:115:VAL:HG21	1.89	0.54
1:X:175:VAL:HG22	1:X:175:VAL:O	2.08	0.54
1:X:338:LEU:C	1:X:338:LEU:HD23	2.27	0.54
1:X:409:ASN:ND2	1:X:410:TYR:H	1.94	0.54
1:X:158:GLN:C	7:X:921:HOH:O	2.45	0.54
1:X:333:ALA:HB2	6:X:804:PG4:H22	1.89	0.54
2:Y:59:TRP:HB2	2:Y:114:PRO:HG2	1.90	0.54
1:X:521:MET:HE3	1:X:525:GLU:HB3	1.88	0.53
1:X:226:TRP:CZ2	1:X:234:SER:HB3	2.44	0.53
1:X:718:THR:O	1:X:721:ILE:HG13	2.09	0.53
1:X:656:ILE:HG23	2:Y:115:VAL:HG22	1.91	0.52
1:X:367:VAL:HG22	1:X:410:TYR:CD2	2.45	0.51
1:X:99:LYS:HA	1:X:99:LYS:HZ2	1.75	0.51
1:X:25:SER:O	1:X:61:GLN:HG3	2.10	0.51
1:X:699:TYR:CG	1:X:723:LYS:HD2	2.46	0.51
1:X:723:LYS:HE2	7:X:822:HOH:O	2.10	0.51
1:X:350:MET:HE2	1:X:352:LEU:HD13	1.92	0.50
1:X:69:ILE:HD11	1:X:116:ILE:HD11	1.92	0.49
1:X:74:SER:HA	1:X:108:LEU:O	2.12	0.49
1:X:17:GLN:HA	1:X:39:THR:HG22	1.94	0.49
1:X:411:TYR:C	1:X:411:TYR:CD1	2.85	0.49
2:Y:107:GLU:HB3	7:Y:419:HOH:O	2.12	0.48
1:X:153:LEU:HB3	1:X:154:PRO:HD2	1.95	0.48
1:X:240:THR:CG2	1:X:242:ARG:HD2	2.43	0.48
1:X:453:ASP:HB2	7:X:903:HOH:O	2.13	0.48
1:X:17:GLN:HE21	1:X:39:THR:HG23	1.79	0.47
1:X:203:LEU:HD23	1:X:203:LEU:C	2.35	0.47
2:Y:191:ASP:HB2	2:Y:195:GLN:HG3	1.96	0.47
1:X:61:GLN:OE1	1:X:125:PRO:HG3	2.14	0.47
1:X:585:ALA:O	1:X:586:LYS:HB2	2.15	0.47
2:Y:191:ASP:HB2	2:Y:195:GLN:CG	2.44	0.47
1:X:327:SER:CB	7:X:915:HOH:O	2.64	0.46
2:Y:206:LEU:HG	2:Y:206:LEU:O	2.15	0.46
1:X:520:ASP:HB3	7:X:914:HOH:O	2.17	0.45
2:Y:163:ALA:O	2:Y:166:SER:HB2	2.17	0.45
1:X:231:ASP:HB2	1:X:232:PRO:CD	2.47	0.45
1:X:699:TYR:CD2	1:X:723:LYS:HD2	2.52	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:X:444:LYS:HA	1:X:709:ASN:ND2	2.32	0.45
1:X:240:THR:CG2	1:X:242:ARG:HH11	2.28	0.44
1:X:593:ASP:OD1	1:X:595:ARG:CG	2.66	0.44
1:X:617:ARG:HG2	7:X:871:HOH:O	2.17	0.44
1:X:69:ILE:HG22	1:X:150:ASN:HB3	2.00	0.44
1:X:274:ASN:OD1	1:X:359:ARG:HB2	2.18	0.44
1:X:424:GLN:HG3	7:X:886:HOH:O	2.16	0.44
1:X:85:ASN:N	1:X:85:ASN:OD1	2.51	0.44
1:X:52:ASN:HB3	1:X:211:HIS:HE2	1.83	0.43
1:X:197:LYS:HB3	1:X:198:ASN:H	1.72	0.43
1:X:54:PRO:HG2	1:X:57:ASN:ND2	2.33	0.43
1:X:178:ARG:NE	1:X:224:GLU:OE2	2.49	0.43
1:X:293:THR:HB	1:X:332:VAL:CG1	2.49	0.43
1:X:360:LEU:C	1:X:360:LEU:HD23	2.40	0.43
1:X:609:GLU:O	1:X:613:LYS:HG3	2.19	0.42
1:X:298:THR:HG21	7:X:821:HOH:O	2.19	0.42
1:X:577:VAL:O	1:X:581:ILE:HG12	2.19	0.42
2:Y:98:ASP:O	2:Y:102:ILE:HD12	2.18	0.42
1:X:367:VAL:HG22	1:X:410:TYR:CE2	2.54	0.42
1:X:69:ILE:HD12	1:X:114:TYR:HB2	2.00	0.42
1:X:388:ASN:O	1:X:388:ASN:CG	2.57	0.42
1:X:460:ALA:HA	1:X:471:VAL:HA	2.01	0.42
1:X:78:THR:O	1:X:136:TRP:HA	2.19	0.42
1:X:345:THR:HG22	1:X:443:GLU:OE1	2.20	0.42
2:Y:199:VAL:CG1	2:Y:200:LYS:N	2.83	0.42
1:X:142:LYS:HE3	1:X:142:LYS:HB3	1.86	0.42
1:X:137:THR:HA	1:X:142:LYS:O	2.20	0.42
1:X:149:ASP:N	1:X:149:ASP:OD1	2.52	0.42
1:X:660:TYR:HB2	1:X:707:ILE:HG12	2.01	0.41
1:X:240:THR:HG21	1:X:242:ARG:NH1	2.31	0.41
2:Y:199:VAL:HG13	2:Y:200:LYS:N	2.35	0.41
2:Y:48:VAL:O	2:Y:145:ALA:HA	2.20	0.41
1:X:656:ILE:CD1	2:Y:115:VAL:HG11	2.47	0.41
2:Y:207:LYS:O	2:Y:207:LYS:HE2	2.21	0.41
1:X:599:ASP:HB2	7:X:876:HOH:O	2.20	0.40
2:Y:49:LEU:HD11	2:Y:66:VAL:HG21	2.02	0.40
2:Y:57:ASN:HB2	7:Y:416:HOH:O	2.21	0.40
1:X:656:ILE:HD12	2:Y:115:VAL:CG2	2.50	0.40
1:X:699:TYR:HB3	1:X:723:LYS:HB2	2.03	0.40
1:X:340:LEU:HD12	1:X:340:LEU:HA	1.89	0.40
1:X:668:LEU:HD12	1:X:668:LEU:HA	1.92	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	X	668/735 (91%)	630 (94%)	34 (5%)	4 (1%)	30 50
2	Y	168/189 (89%)	160 (95%)	8 (5%)	0	100 100
All	All	836/924 (90%)	790 (94%)	42 (5%)	4 (0%)	34 55

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	102	ASN
1	X	377	VAL
1	X	125	PRO
1	X	124	ASN

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	X	609/661 (92%)	520 (85%)	89 (15%)	4 7
2	Y	142/156 (91%)	120 (84%)	22 (16%)	3 6
All	All	751/817 (92%)	640 (85%)	111 (15%)	4 7

All (111) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	17	GLN
1	X	28	ASN
1	X	45	ILE
1	X	47	SER
1	X	55	SER
1	X	66	SER
1	X	69	ILE
1	X	70	LYS
1	X	74	SER
1	X	76	GLU
1	X	93	ASP
1	X	96	VAL
1	X	99	LYS
1	X	102	ASN
1	X	105	LYS
1	X	106	ILE
1	X	107	ARG
1	X	108	LEU
1	X	110	LYS
1	X	112	ARG
1	X	115	GLN
1	X	117	LYS
1	X	122	ARG
1	X	124	ASN
1	X	127	GLU
1	X	133	LYS
1	X	136	TRP
1	X	140	GLN
1	X	149	ASP
1	X	151	LEU
1	X	197	LYS
1	X	199	LYS
1	X	207	ILE
1	X	209	ASN
1	X	212	GLU
1	X	214	LYS
1	X	216	LEU
1	X	240	THR
1	X	245	LYS
1	X	260	PRO
1	X	275	GLU
1	X	288	THR
1	X	298	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	X	302	GLU
1	X	321	SER
1	X	325	SER
1	X	338	LEU
1	X	340	LEU
1	X	344	ARG
1	X	345	THR
1	X	348	GLU
1	X	357	THR
1	X	388	ASN
1	X	401	LEU
1	X	404	ILE
1	X	409	ASN
1	X	413	SER
1	X	424	GLN
1	X	428	SER
1	X	437	ASN
1	X	444	LYS
1	X	447	GLN
1	X	465	GLU
1	X	466	ASN
1	X	468	ARG
1	X	470	ARG
1	X	486	GLU
1	X	509	ASN
1	X	515	GLU
1	X	518	LYS
1	X	524	LYS
1	X	527	LEU
1	X	556	GLN
1	X	557	GLN
1	X	563	LYS
1	X	576	THR
1	X	595	ARG
1	X	629	LEU
1	X	633	LYS
1	X	649	THR
1	X	654	GLU
1	X	666	SER
1	X	707	ILE
1	X	711	SER
1	X	712	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	X	713	ASN
1	X	722	LYS
1	X	723	LYS
1	X	730	LYS
2	Y	51	LYS
2	Y	68	GLN
2	Y	69	LEU
2	Y	72	ARG
2	Y	74	VAL
2	Y	80	LEU
2	Y	99	ARG
2	Y	110	LYS
2	Y	111	ARG
2	Y	113	SER
2	Y	133	LYS
2	Y	137	LEU
2	Y	138	LYS
2	Y	140	SER
2	Y	160	GLU
2	Y	166	SER
2	Y	168	SER
2	Y	172	SER
2	Y	194	GLU
2	Y	195	GLN
2	Y	204	GLN
2	Y	207	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (23) such sidechains are listed below:

Mol	Chain	Res	Type
1	X	17	GLN
1	X	28	ASN
1	X	102	ASN
1	X	104	ASN
1	X	140	GLN
1	X	158	GLN
1	X	180	ASN
1	X	209	ASN
1	X	263	HIS
1	X	268	ASN
1	X	328	ASN
1	X	363	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	X	400	GLN
1	X	403	GLN
1	X	409	ASN
1	X	466	ASN
1	X	483	GLN
1	X	509	ASN
1	X	564	ASN
1	X	584	ASN
1	X	601	ASN
2	Y	64	ASN
2	Y	88	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 5 are monoatomic - leaving 1 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	PG4	X	804	-	12,12,12	0.67	0	11,11,11	0.92	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	PG4	X	804	-	-	0/10/10/10	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	X	804	PG4	1	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^{th} 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^2)	Q<0.9
1	X	676/735 (91%)	-0.15	23 (3%) 49 54	5, 20, 40, 49	0
2	Y	170/189 (89%)	0.06	5 (2%) 55 60	14, 23, 40, 54	0
All	All	846/924 (91%)	-0.10	28 (3%) 50 55	5, 21, 40, 54	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	Y	41	ARG	5.3
1	X	124	ASN	4.5
1	X	735	GLY	3.9
1	X	213	LYS	3.9
1	X	303	VAL	3.8
1	X	100	ALA	3.8
2	Y	42	ALA	3.3
1	X	209	ASN	3.2
2	Y	74	VAL	3.0
1	X	85	ASN	2.9
1	X	713	ASN	2.8
1	X	101	SER	2.7
1	X	670	GLN	2.7
1	X	102	ASN	2.6
1	X	94	GLN	2.6
1	X	301	SER	2.5
1	X	76	GLU	2.4
1	X	112	ARG	2.4
1	X	199	LYS	2.4
1	X	212	GLU	2.3
1	X	127	GLU	2.3
1	X	302	GLU	2.2
1	X	149	ASP	2.2
1	X	425	ASP	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	Y	208	GLY	2.1
2	Y	77	GLU	2.1
1	X	71	VAL	2.0
1	X	600	ARG	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^{th} 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^2)	Q<0.9
6	PG4	X	804	13/13	0.94	0.17	2.67	27,32,37,42	0
5	NA	X	803	1/1	0.97	0.14	2.04	28,28,28,28	0
3	CA	X	800	1/1	0.99	0.10	-0.80	29,29,29,29	0
4	MN	Y	303	1/1	0.98	0.09	-1.48	33,33,33,33	0
3	CA	X	801	1/1	1.00	0.09	-1.70	26,26,26,26	0
5	NA	X	802	1/1	0.98	0.08	-2.18	28,28,28,28	0

6.5 Other polymers (i)

There are no such residues in this entry.