Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 02:00 PM GMT

PDB ID	:	3VS1
Title	:	Crystal structure of HCK complexed with a pyrrolo-pyrimidine inhibitor 1-[4-(4-amino-7-cyclopentyl-7H-pyrrolo[2,3-d]pyrimidin-5-yl)phenyl]-3-phenylurea
Authors	:	Kuratani, M.; Tomabechi, Y.; Toyama, M.; Handa, N.; Yokoyama, S.
Deposited on	:	2012-04-21
Resolution	:	2.46 Å(reported)

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

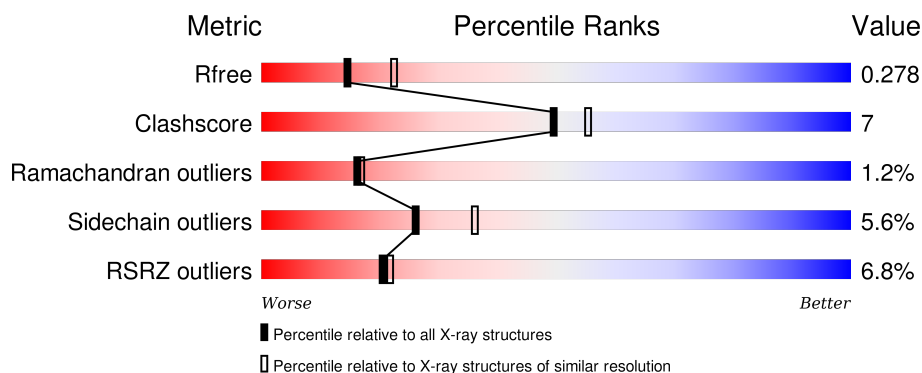|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA,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

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
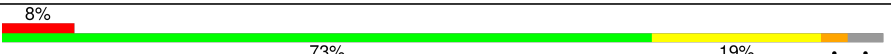*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.46 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	Percentile Ranks	Value
Rfree		0.278
Clashscore		7
Ramachandran outliers		1.2%
Sidechain outliers		5.6%
RSRZ outliers		6.8%

Worse ← → *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	91344	4776 (2.50-2.42)
Clashscore	102246	1030 (2.48-2.44)
Ramachandran outliers	100387	1024 (2.48-2.44)
Sidechain outliers	100360	1024 (2.48-2.44)
RSRZ outliers	91569	4787 (2.50-2.42)

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	A	454	4% / 77% / 15% / · 6%
1	B	454	8% / 73% / 19% / · ·

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7069 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 Tyrosine-protein kinase HCK.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	425	Total	C	N	O	P	S	0	0	0
			3426	2187	578	640	1	20			
1	B	434	Total	C	N	O	P	S	0	0	0
			3496	2233	588	654	1	20			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	78	GLY	-	EXPRESSION TAG	UNP P08631
A	79	ALA	-	EXPRESSION TAG	UNP P08631
A	80	MET	-	EXPRESSION TAG	UNP P08631
A	81	GLY	-	EXPRESSION TAG	UNP P08631
A	82	SER	-	EXPRESSION TAG	UNP P08631
A	83	GLY	-	EXPRESSION TAG	UNP P08631
A	84	ILE	-	EXPRESSION TAG	UNP P08631
A	85	ARG	-	EXPRESSION TAG	UNP P08631
A	528	GLU	GLN	ENGINEERED MUTATION	UNP P08631
A	529	GLU	GLN	ENGINEERED MUTATION	UNP P08631
A	530	ILE	GLN	ENGINEERED MUTATION	UNP P08631
B	78	GLY	-	EXPRESSION TAG	UNP P08631
B	79	ALA	-	EXPRESSION TAG	UNP P08631
B	80	MET	-	EXPRESSION TAG	UNP P08631
B	81	GLY	-	EXPRESSION TAG	UNP P08631
B	82	SER	-	EXPRESSION TAG	UNP P08631
B	83	GLY	-	EXPRESSION TAG	UNP P08631
B	84	ILE	-	EXPRESSION TAG	UNP P08631
B	85	ARG	-	EXPRESSION TAG	UNP P08631
B	528	GLU	GLN	ENGINEERED MUTATION	UNP P08631
B	529	GLU	GLN	ENGINEERED MUTATION	UNP P08631
B	530	ILE	GLN	ENGINEERED MUTATION	UNP P08631

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total Ca 1 1		0	0
2	A	1	Total Ca 1 1		0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total Cl 1 1		0	0
3	A	1	Total Cl 1 1		0	0

- Molecule 4 is 1-[4-(4-AMINO-7-CYCLOPENTYL-7H-PYRROLO[2,3-D]PYRIMIDIN-5-YL)PHENYL]-3-PHENYLUREA (three-letter code: VSA) (formula: $C_{24}H_{24}N_6O$).
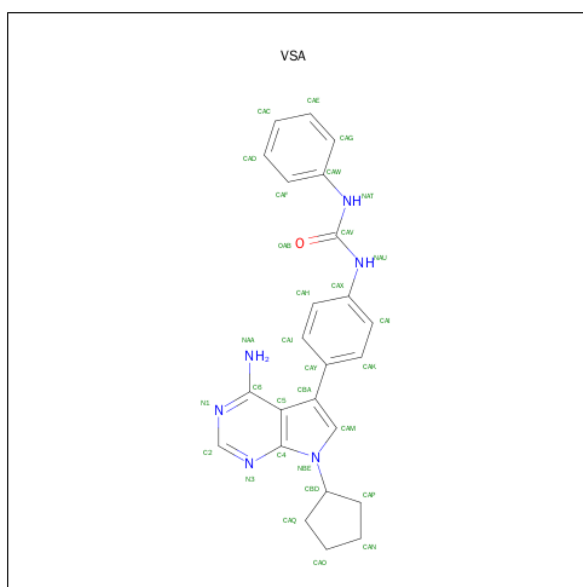
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 31	C 24	N 6	O 1	0	0
4	B	1	Total 31	C 24	N 6	O 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	52	Total O 52 52		0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	29	Total 29	O 29	0	0

3 Residue-property plots ⓘ

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: Tyrosine-protein kinase HCK

Chain A:

- Molecule 1: Tyrosine-protein kinase HCK

Chain B:
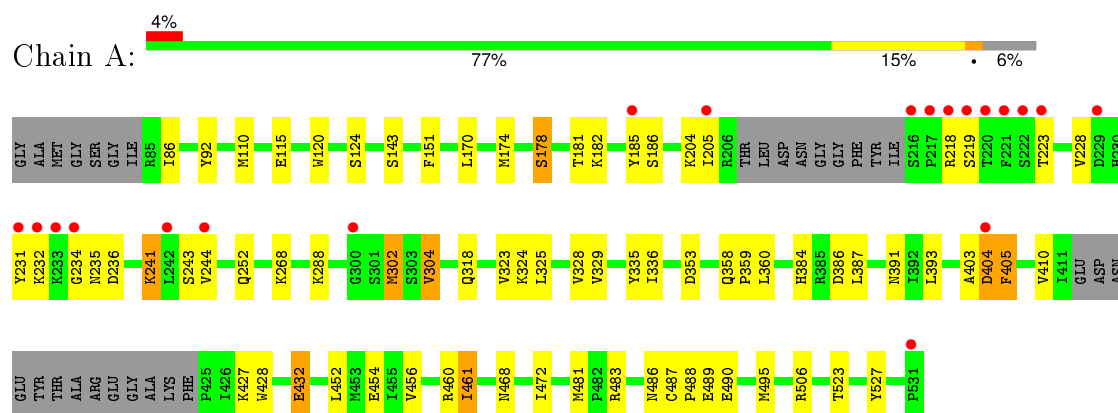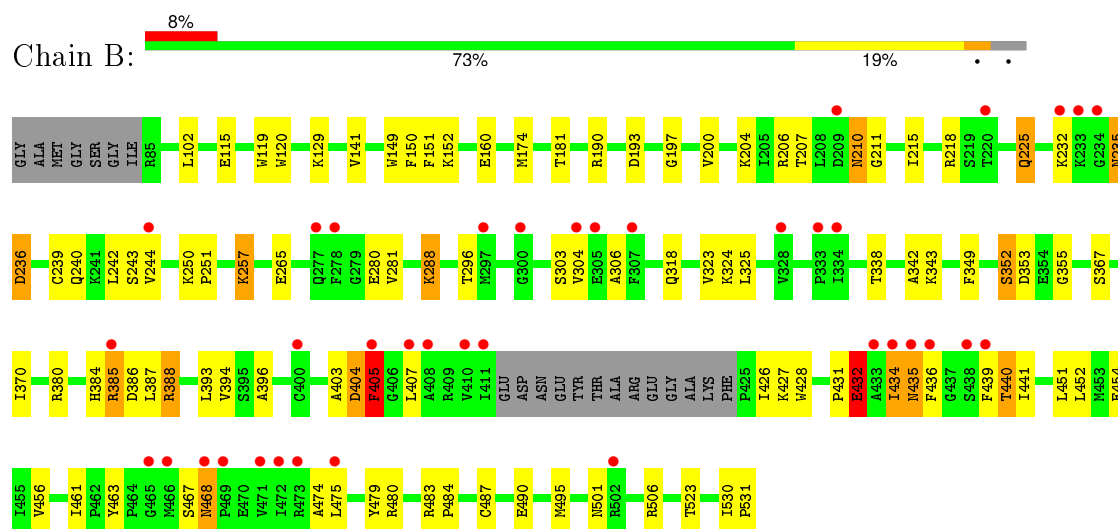
4   Data and refinement statistics   (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	49.03Å 73.27Å 179.85Å 90.00° 96.21° 90.00°	Depositor
Resolution (Å)	44.70 – 2.46 44.70 – 2.46	Depositor EDS
% Data completeness (in resolution range)	99.5 (44.70-2.46) 99.6 (44.70-2.46)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.42 (at 2.45Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.240 , 0.281 0.235 , 0.278	Depositor DCC
R_{free} test set	2296 reflections (5.27%)	DCC
Wilson B-factor (Å^2)	54.0	Xtriage
Anisotropy	0.597	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 47.8	EDS
Estimated twinning fraction	0.031 for h,-k,-h-l	Xtriage
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	1 of 45904 reflections (0.002%)	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7069	wwPDB-VP
Average B, all atoms (Å^2)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.57% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: VSA, CA, PTR, CL

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	0.25	0/3489	0.45	0/4707
1	B	0.24	0/3562	0.46	1/4808 (0.0%)
All	All	0.24	0/7051	0.46	1/9515 (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	B	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	405	PHE	CB-CG-CD1	5.05	124.33	120.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	404	ASP	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	3426	0	3391	41	0
1	B	3496	0	3455	58	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	31	0	24	2	0
4	B	31	0	24	5	0
5	A	52	0	0	1	0
5	B	29	0	0	1	0
All	All	7069	0	6894	99	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 7.

All (99) 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:404:ASP:HA	1:A:405:PHE:HB3	1.60	0.84
1:B:404:ASP:HA	1:B:405:PHE:HB3	1.62	0.82
1:A:318:GLN:OE1	1:A:324:LYS:NZ	2.24	0.70
1:B:149:TRP:HZ2	1:B:225:GLN:HG2	1.57	0.69
1:B:428:TRP:NE1	1:B:454:GLU:OE1	2.23	0.69
1:A:384:HIS:CD2	1:A:405:PHE:HB2	2.28	0.68
1:A:432:GLU:OE2	1:A:506:ARG:NH1	2.21	0.67
1:B:483:ARG:NH1	1:B:487:CYS:O	2.27	0.67
1:B:257:LYS:HD3	1:B:257:LYS:H	1.59	0.66
1:A:181:THR:OG1	1:A:204:LYS:NZ	2.29	0.66
4:B:602:VSA:H14	4:B:602:VSA:H13	1.60	0.65
1:B:318:GLN:OE1	1:B:324:LYS:NZ	2.29	0.65
1:A:325:LEU:HD21	4:A:602:VSA:H20	1.80	0.64
1:B:384:HIS:CD2	1:B:405:PHE:HB2	2.35	0.62
1:B:303:SER:HB3	1:B:306:ALA:HB3	1.81	0.61
1:B:215:ILE:HG22	1:B:530:ILE:HG21	1.82	0.61
1:A:232:LYS:O	1:A:241:LYS:NZ	2.30	0.61
1:A:252:GLN:NE2	5:A:747:HOH:O	2.32	0.60
1:B:452:LEU:HD23	1:B:495:MET:HG2	1.83	0.59
1:A:268:LYS:HE3	1:A:288:LYS:HE3	1.84	0.59
1:A:243:SER:OG	1:A:244:VAL:N	2.36	0.59
1:B:210:ASN:N	1:B:210:ASN:OD1	2.34	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:218:ARG:NH1	1:B:236:ASP:OD2	2.35	0.59
1:A:428:TRP:NE1	1:A:454:GLU:OE1	2.35	0.59
1:B:480:ARG:NH2	1:B:501:ASN:OD1	2.36	0.58
1:B:243:SER:OG	1:B:244:VAL:N	2.37	0.58
1:B:432:GLU:OE2	1:B:506:ARG:NH1	2.23	0.57
1:B:352:SER:OG	1:B:353:ASP:N	2.36	0.57
1:A:483:ARG:NH1	1:A:487:CYS:O	2.36	0.57
1:B:435:ASN:OD1	1:B:435:ASN:N	2.33	0.57
1:B:265:GLU:N	1:B:265:GLU:OE1	2.33	0.56
1:A:452:LEU:HD23	1:A:495:MET:HG2	1.88	0.56
1:A:393:LEU:HG	1:A:403:ALA:HB2	1.88	0.55
1:A:456:VAL:HG11	1:A:487:CYS:HB2	1.89	0.55
1:B:370:ILE:HD13	1:B:451:LEU:HD21	1.89	0.55
1:B:102:LEU:HD21	1:B:129:LYS:HB2	1.87	0.55
4:B:602:VSA:CAK	4:B:602:VSA:H13	2.21	0.54
1:B:474:ALA:HB1	1:B:479:TYR:HB3	1.91	0.53
1:B:181:THR:OG1	1:B:204:LYS:NZ	2.39	0.53
1:B:456:VAL:HG11	1:B:487:CYS:HB2	1.92	0.52
1:A:235:ASN:N	1:A:235:ASN:OD1	2.42	0.52
1:A:323:VAL:HG21	1:A:393:LEU:HD12	1.91	0.51
1:A:234:GLY:H	1:A:241:LYS:HD2	1.75	0.51
1:B:119:TRP:CD1	1:B:257:LYS:HG3	2.45	0.51
1:A:386:ASP:OD2	1:A:391:ASN:ND2	2.44	0.51
1:B:432:GLU:HA	1:B:436:PHE:HD2	1.76	0.50
1:B:303:SER:OG	1:B:304:VAL:N	2.43	0.50
1:B:235:ASN:ND2	1:B:235:ASN:O	2.44	0.50
1:B:349:PHE:O	1:B:355:GLY:HA3	2.12	0.50
1:A:523:THR:O	1:B:490:GLU:HG2	2.11	0.50
1:B:427:LYS:HD3	1:B:463:TYR:HD1	1.77	0.49
1:B:386:ASP:OD1	1:B:388:ARG:NE	2.46	0.48
1:A:110:MET:HB3	1:A:124:SER:HA	1.93	0.48
1:B:151:PHE:CD1	1:B:174:MET:HB2	2.48	0.48
1:B:115:GLU:HA	1:B:120:TRP:CD1	2.49	0.47
1:A:490:GLU:HG2	1:B:523:THR:O	2.14	0.47
1:B:210:ASN:N	1:B:211:GLY:HA2	2.28	0.47
1:B:160:GLU:HA	1:B:200:VAL:HG21	1.96	0.47
1:B:232:LYS:HA	1:B:242:LEU:HB2	1.96	0.47
1:A:288:LYS:N	1:A:288:LYS:HD2	2.30	0.46
1:A:489:GLU:OE2	1:B:206:ARG:NH2	2.49	0.46
1:A:178:SER:OG	1:A:527:PTR:O3P	2.30	0.46
1:A:329:VAL:HB	1:A:335:TYR:HB2	1.97	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:178:SER:OG	1:A:186:SER:OG	2.35	0.45
1:B:338:THR:HG21	4:B:602:VSA:H15	1.98	0.45
1:B:530:ILE:HD12	1:B:531:PRO:HD2	1.98	0.45
1:A:360:LEU:HD21	1:A:488:PRO:HD3	1.99	0.44
1:A:228:VAL:O	1:A:232:LYS:HG3	2.16	0.44
1:B:349:PHE:O	1:B:352:SER:HB3	2.17	0.44
1:B:193:ASP:O	1:B:197:GLY:N	2.43	0.43
1:A:231:TYR:CE1	1:A:236:ASP:HB3	2.52	0.43
1:B:150:PHE:HE2	1:B:152:LYS:HD2	1.83	0.43
1:B:440:THR:OG1	1:B:441:ILE:N	2.51	0.43
1:B:393:LEU:HG	1:B:403:ALA:HB2	2.01	0.43
1:A:185:TYR:O	1:A:205:ILE:HB	2.17	0.43
1:A:487:CYS:HA	1:A:488:PRO:HD3	1.89	0.43
1:A:468:ASN:O	1:A:472:ILE:HG13	2.19	0.43
1:B:385:ARG:HD2	1:B:439:PHE:CG	2.54	0.43
1:B:467:SER:OG	1:B:468:ASN:N	2.52	0.42
1:A:328:VAL:HG23	1:A:336:ILE:HD13	2.01	0.42
1:A:460:ARG:NH2	1:A:481:MET:SD	2.92	0.42
1:A:151:PHE:CD1	1:A:174:MET:HB2	2.55	0.42
1:B:343:LYS:HB2	1:B:394:VAL:HB	2.02	0.42
1:A:432:GLU:HG3	1:A:432:GLU:H	1.39	0.42
1:B:250:LYS:HA	1:B:251:PRO:HD2	1.94	0.42
1:B:323:VAL:CG1	1:B:403:ALA:HA	2.51	0.41
1:A:115:GLU:HA	1:A:120:TRP:CD1	2.56	0.41
1:B:239:CYS:SG	1:B:240:GLN:N	2.94	0.41
1:B:288:LYS:NZ	5:B:705:HOH:O	2.53	0.41
1:B:530:ILE:HA	1:B:531:PRO:HD2	1.97	0.41
1:B:483:ARG:HA	1:B:484:PRO:HD3	1.87	0.41
1:B:119:TRP:CG	1:B:257:LYS:HG3	2.56	0.41
1:A:325:LEU:HD21	4:A:602:VSA:CAF	2.49	0.41
1:B:342:ALA:HB3	1:B:396:ALA:HB2	2.02	0.41
1:A:358:GLN:HA	1:A:359:PRO:HD3	1.89	0.41
1:A:427:LYS:NZ	1:A:461:ILE:HD11	2.36	0.41
1:B:380:ARG:HD3	1:B:380:ARG:HA	1.80	0.41
4:B:602:VSA:H14	4:B:602:VSA:NAA	2.32	0.40
1:B:281:VAL:HG11	4:B:602:VSA:CAM	2.51	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	418/454 (92%)	392 (94%)	20 (5%)	6 (1%)	14	13
1	B	429/454 (94%)	402 (94%)	23 (5%)	4 (1%)	21	25
All	All	847/908 (93%)	794 (94%)	43 (5%)	10 (1%)	16	17

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	182	LYS
1	A	405	PHE
1	A	302	MET
1	B	405	PHE
1	B	432	GLU
1	A	304	VAL
1	A	486	ASN
1	B	434	ILE
1	B	431	PRO
1	A	410	VAL

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	370/393 (94%)	354 (96%)	16 (4%)	35	50
1	B	377/393 (96%)	351 (93%)	26 (7%)	19	25
All	All	747/786 (95%)	705 (94%)	42 (6%)	26	36

All (42) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	86	ILE
1	A	92	TYR
1	A	143	SER
1	A	170	LEU
1	A	178	SER
1	A	218	ARG
1	A	219	SER
1	A	223	THR
1	A	241	LYS
1	A	302	MET
1	A	304	VAL
1	A	353	ASP
1	A	387	LEU
1	A	404	ASP
1	A	432	GLU
1	A	461	ILE
1	B	141	VAL
1	B	190	ARG
1	B	207	THR
1	B	210	ASN
1	B	225	GLN
1	B	235	ASN
1	B	236	ASP
1	B	257	LYS
1	B	280	GLU
1	B	288	LYS
1	B	296	THR
1	B	325	LEU
1	B	352	SER
1	B	367	SER
1	B	385	ARG
1	B	387	LEU
1	B	388	ARG
1	B	407	LEU
1	B	426	ILE
1	B	432	GLU
1	B	434	ILE
1	B	435	ASN
1	B	440	THR
1	B	461	ILE
1	B	468	ASN
1	B	475	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 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	#$|Z| > 2$
1	PTR	A	527	1,2	14,16,17	1.30	1 (7%)	18,22,24	0.68	0
1	PTR	B	527	1,2	14,16,17	1.29	1 (7%)	18,22,24	0.75	1 (5%)

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
1	PTR	A	527	1,2	-	0/9/11/13	0/1/1/1
1	PTR	B	527	1,2	-	0/9/11/13	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	527	PTR	OH-CZ	-4.29	1.30	1.40
1	B	527	PTR	OH-CZ	-4.28	1.30	1.40

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	527	PTR	O-C-CA	-2.23	119.68	125.49

Full wwPDB X-ray Structure Validation Report

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
1	A	527	PTR	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	VSA	A	602	-	33,35,35	0.91	3 (9%)	36,49,49	2.01	5 (13%)
4	VSA	B	602	-	33,35,35	0.91	3 (9%)	36,49,49	2.11	6 (16%)

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	VSA	A	602	-	-	0/12/23/23	0/5/5/5
4	VSA	B	602	-	-	0/12/23/23	0/5/5/5

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	602	VSA	CAM-NBE	-3.06	1.33	1.38

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	602	VSA	CAM-NBE	-3.05	1.33	1.38
4	B	602	VSA	CAV-NAU	-2.26	1.33	1.37
4	B	602	VSA	CAV-NAT	-2.19	1.33	1.37
4	A	602	VSA	CAV-NAU	-2.15	1.33	1.37
4	A	602	VSA	CAV-NAT	-2.07	1.33	1.37

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	602	VSA	N3-C2-N1	-9.77	121.41	128.89
4	A	602	VSA	N3-C2-N1	-9.77	121.41	128.89
4	B	602	VSA	CAJ-CAY-CBA	-3.61	115.02	120.87
4	A	602	VSA	CAJ-CAY-CBA	-2.79	116.35	120.87
4	B	602	VSA	CAM-CBA-CAY	-2.69	120.25	125.37
4	A	602	VSA	CAM-CBA-CAY	-2.21	121.17	125.37
4	B	602	VSA	CAX-NAU-CAV	2.04	130.59	126.65
4	A	602	VSA	CAQ-CBD-CAP	2.04	106.44	104.32
4	A	602	VSA	CAX-NAU-CAV	2.16	130.82	126.65
4	B	602	VSA	CAQ-CBD-CAP	2.47	106.87	104.32
4	B	602	VSA	CAK-CAY-CBA	2.70	125.23	120.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	602	VSA	2	0
4	B	602	VSA	5	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 (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, 95th 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(Å²)	Q<0.9
1	A	424/454 (93%)	0.52	20 (4%)	35	38	37, 63, 103, 117	1 (0%)
1	B	433/454 (95%)	0.73	38 (8%)	12	13	42, 74, 116, 129	1 (0%)
All	All	857/908 (94%)	0.63	58 (6%)	20	22	37, 69, 112, 129	2 (0%)

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	475	LEU	8.9
1	B	436	PHE	6.1
1	B	468	ASN	5.4
1	A	220	THR	5.2
1	A	234	GLY	5.1
1	B	304	VAL	5.1
1	B	469	PRO	4.7
1	B	439	PHE	4.0
1	B	433	ALA	4.0
1	B	297	MET	4.0
1	B	434	ILE	3.9
1	B	435	ASN	3.8
1	A	233	LYS	3.6
1	B	471	VAL	3.6
1	B	473	ARG	3.6
1	B	278	PHE	3.6
1	A	222	SER	3.5
1	B	472	ILE	3.5
1	B	465	GLY	3.3
1	B	407	LEU	3.2
1	A	217	PRO	3.2
1	B	220	THR	3.1
1	B	277	GLN	3.1
1	B	438	SER	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	307	PHE	3.0
1	A	531	PRO	3.0
1	A	205	ILE	3.0
1	A	219	SER	2.9
1	A	223	THR	2.9
1	A	242	LEU	2.9
1	B	410	VAL	2.8
1	B	233	LYS	2.7
1	B	209	ASP	2.7
1	A	221	PHE	2.7
1	B	405	PHE	2.7
1	B	466	MET	2.7
1	B	300	GLY	2.7
1	B	411	ILE	2.6
1	B	232	LYS	2.6
1	B	328	VAL	2.6
1	A	244	VAL	2.6
1	B	333	PRO	2.5
1	B	502	ARG	2.5
1	B	385	ARG	2.5
1	B	408	ALA	2.4
1	B	400	CYS	2.4
1	B	305	GLU	2.4
1	A	216	SER	2.3
1	A	229	ASP	2.3
1	B	334	ILE	2.3
1	A	218	ARG	2.3
1	B	234	GLY	2.3
1	A	231	TYR	2.2
1	A	185	TYR	2.1
1	A	404	ASP	2.0
1	A	232	LYS	2.0
1	A	300	GLY	2.0
1	B	244	VAL	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. 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, 95th 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
1	PTR	B	527	16/17	0.97	0.19	-	47,53,60,60	0
1	PTR	A	527	16/17	0.95	0.13	-	59,65,72,73	0

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, 95th 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
4	VSA	A	602	31/31	0.96	0.20	0.20	39,50,76,86	0
4	VSA	B	602	31/31	0.93	0.21	0.05	48,64,95,97	0
3	CL	B	603	1/1	0.95	0.17	-1.22	78,78,78,78	0
3	CL	A	603	1/1	0.93	0.13	-1.51	66,66,66,66	0
2	CA	A	601	1/1	0.19	0.33	-	88,88,88,88	0
2	CA	B	601	1/1	0.71	0.28	-	65,65,65,65	0

6.5 Other polymers (i)

There are no such residues in this entry.