wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 04:54 AM GMT

PDB ID	:	2ONL
Title	:	Crystal Structure of the p38a-MAPKAP kinase 2 Heterodimer
Authors	:	Ter Haar, E.
Deposited on	:	2007-01-24
Resolution	:	4.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
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:

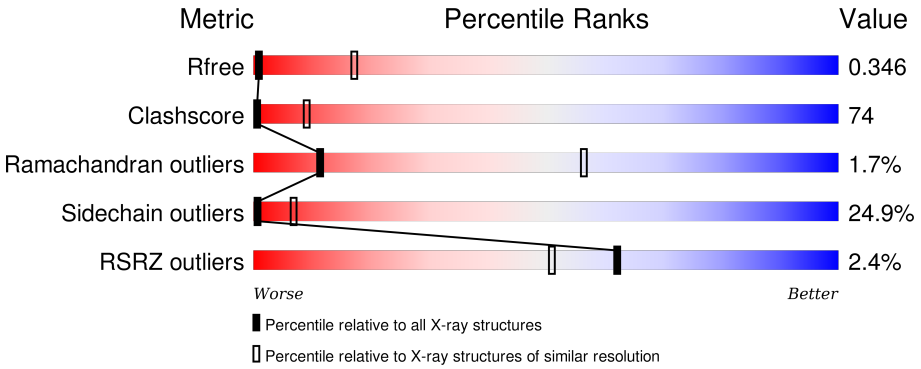
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 4.00 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	1010 (4.42-3.56)
Clashscore	102246	1052 (4.40-3.60)
Ramachandran outliers	100387	1005 (4.40-3.60)
Sidechain outliers	100360	1013 (4.42-3.58)
RSRZ outliers	91569	1013 (4.42-3.56)

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	366	2% 36% 45% 11% 8%
1	B	366	% 33% 48% 11% 7%
2	C	406	3% 10% 43% 23% • 22%
2	D	406	% 14% 42% 19% • 23%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10466 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 Mitogen-activated protein kinase 14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	338	Total	C	N	O	S	0	0	0
			2709	1737	466	494	12			
1	B	339	Total	C	N	O	S	0	0	0
			2714	1740	467	495	12			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	CLONING ARTIFACT	UNP Q16539
A	-4	SER	-	CLONING ARTIFACT	UNP Q16539
A	-3	HIS	-	CLONING ARTIFACT	UNP Q16539
A	-2	MET	-	CLONING ARTIFACT	UNP Q16539
A	-1	LEU	-	CLONING ARTIFACT	UNP Q16539
A	0	GLU	-	CLONING ARTIFACT	UNP Q16539
A	1	MET	-	CLONING ARTIFACT	UNP Q16539
B	-5	GLY	-	CLONING ARTIFACT	UNP Q16539
B	-4	SER	-	CLONING ARTIFACT	UNP Q16539
B	-3	HIS	-	CLONING ARTIFACT	UNP Q16539
B	-2	MET	-	CLONING ARTIFACT	UNP Q16539
B	-1	LEU	-	CLONING ARTIFACT	UNP Q16539
B	0	GLU	-	CLONING ARTIFACT	UNP Q16539
B	1	MET	-	CLONING ARTIFACT	UNP Q16539

- Molecule 2 is a protein called MAP kinase-activated protein kinase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	317	Total	C	N	O	S	0	0	0
			2531	1611	440	462	18			
2	D	313	Total	C	N	O	S	0	0	0
			2512	1597	437	460	18			

There are 12 discrepancies between the modelled and reference sequences:

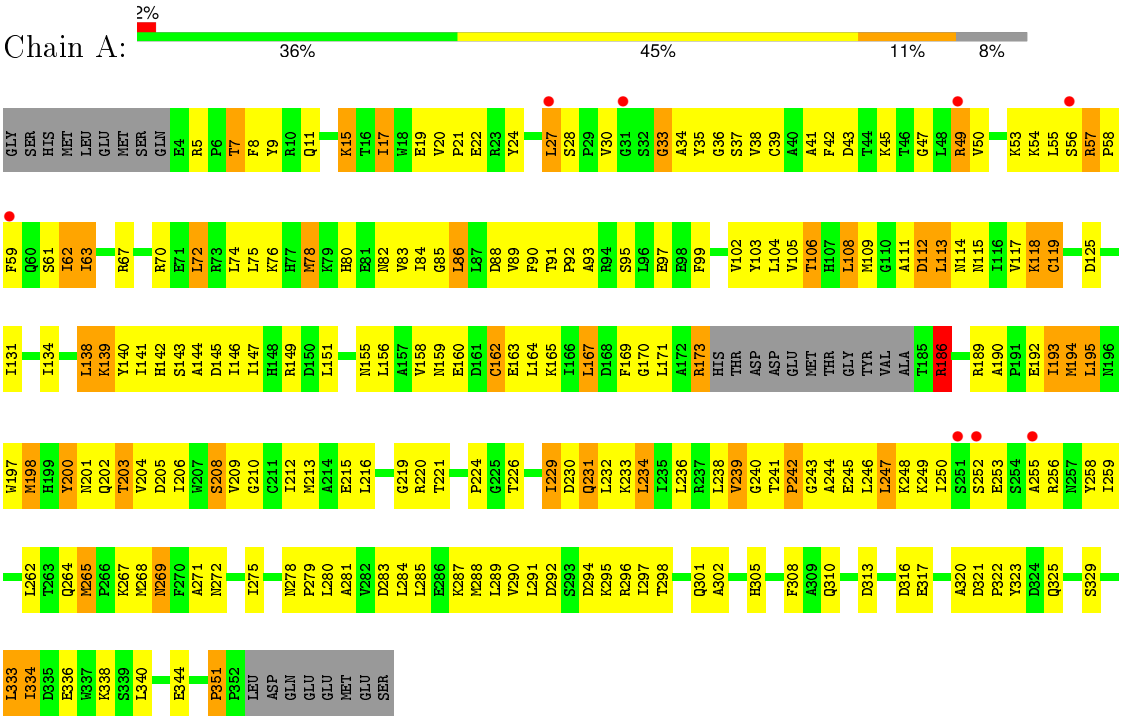
Chain	Residue	Modelled	Actual	Comment	Reference
C	-5	GLY	-	CLONING ARTIFACT	UNP P49137
C	-4	SER	-	CLONING ARTIFACT	UNP P49137
C	-3	HIS	-	CLONING ARTIFACT	UNP P49137
C	-2	MET	-	CLONING ARTIFACT	UNP P49137
C	-1	LEU	-	CLONING ARTIFACT	UNP P49137
C	0	GLU	-	CLONING ARTIFACT	UNP P49137
D	-5	GLY	-	CLONING ARTIFACT	UNP P49137
D	-4	SER	-	CLONING ARTIFACT	UNP P49137
D	-3	HIS	-	CLONING ARTIFACT	UNP P49137
D	-2	MET	-	CLONING ARTIFACT	UNP P49137
D	-1	LEU	-	CLONING ARTIFACT	UNP P49137
D	0	GLU	-	CLONING ARTIFACT	UNP P49137

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.

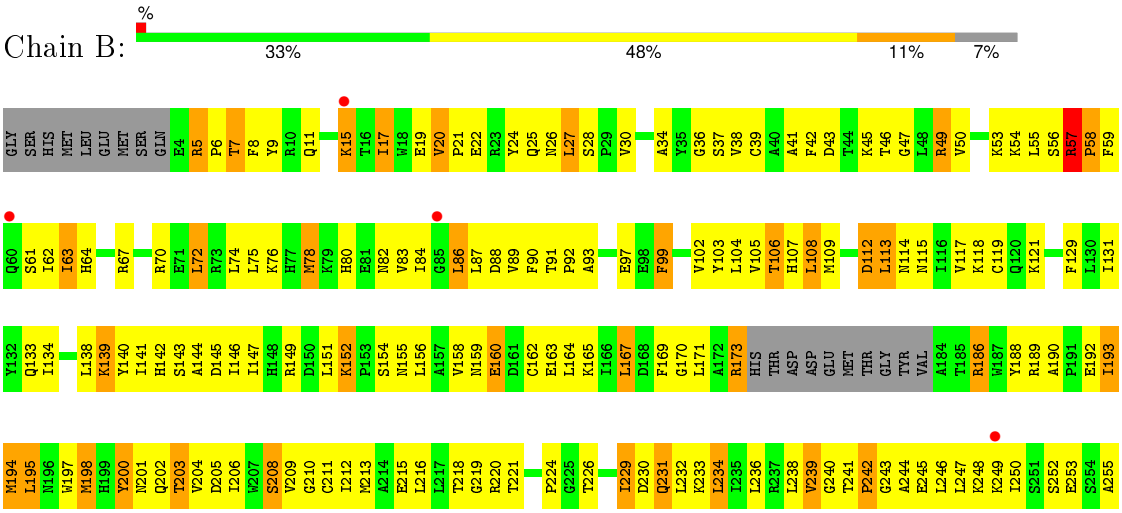
- Molecule 1: Mitogen-activated protein kinase 14

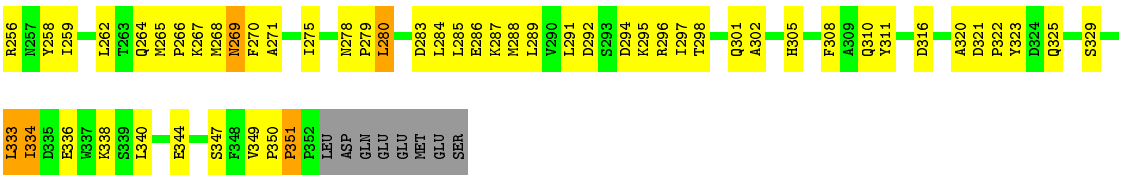
Chain A:

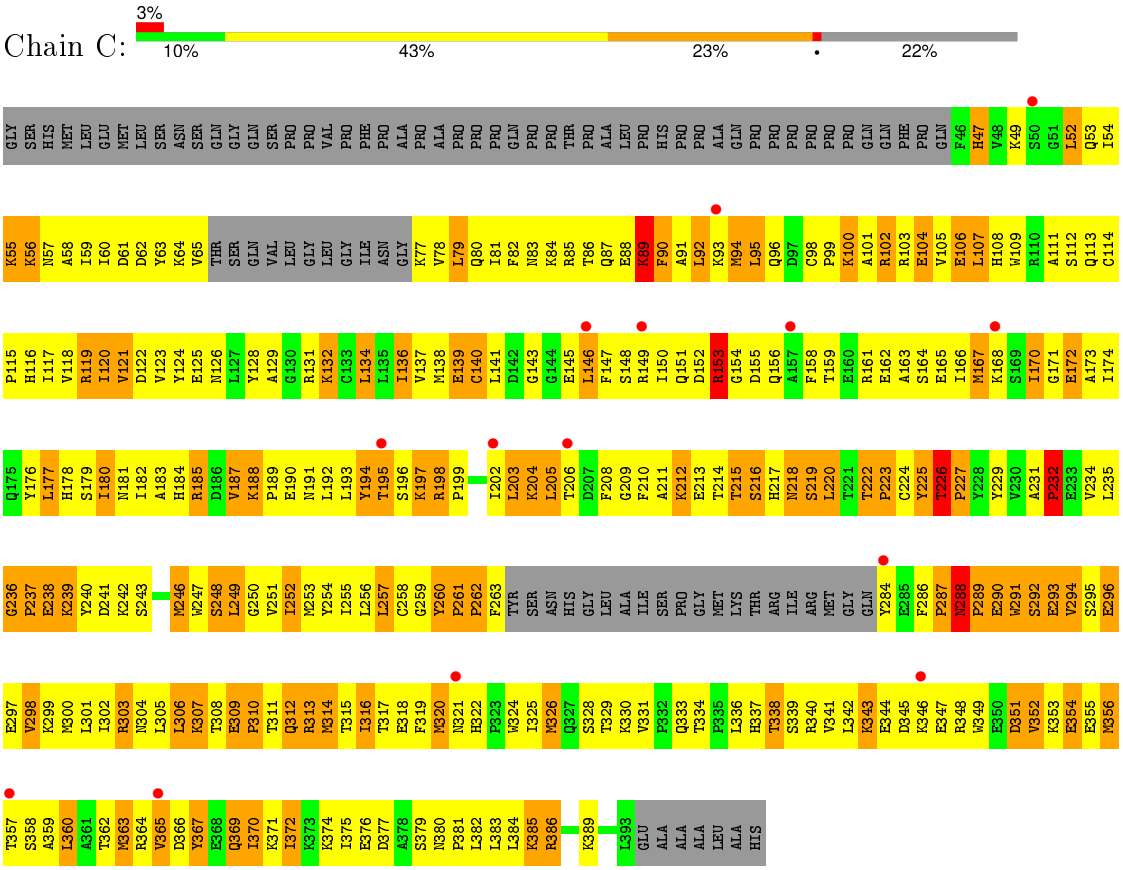


- Molecule 1: Mitogen-activated protein kinase 14

Chain B:





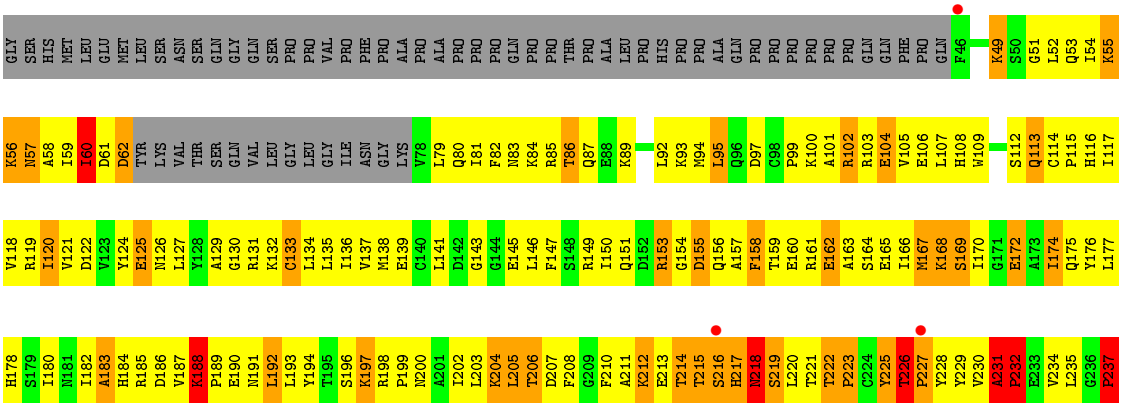


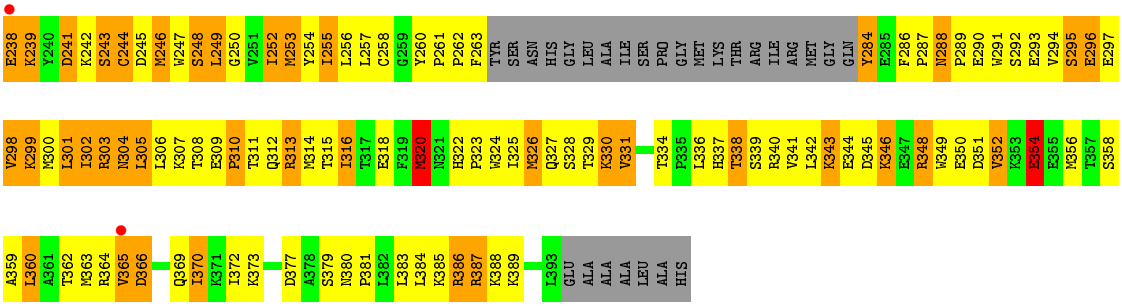
• Molecule 2: MAP kinase-activated protein kinase 2

Chain C:

• Molecule 2: MAP kinase-activated protein kinase 2

Chain D:




E238 K239 Y240 D241 K242 S243 C244 D245 M246 W247 S248 L249 G250 V251 I252 M253 Y254 I255 L256 L257 C258 G259 Y260 P261 P262 F263 TYR SER ASN HIS GLY LEU ALA ILE SER PRO GLY MET LYS THR ARG ILE ARG MET GLY GLN Y284 E285 F286 P287 N288 P289 E290 W291 S292 E293 V294 S295 E296 E297
V298 K299 M300 L301 I302 R303 N304 L305 L306 K307 T308 E309 P310 T311 Q312 R313 M314 T315 I316 T317 E318 F319 M320 N321 H322 P323 W324 I325 M326 Q327 S328 T329 K330 V331 T334 P335 L336 H337 T338 S339 R340 V341 L342 K343 E344 D345 K346 E347 R348 W349 E350 D351 V352 K353 E354 E355 M356 T357 S358
A359 L360 A361 T362 M363 R364 V365 D366 Q369 I370 K371 I372 K373 D377 A378 S379 N380 P381 L382 I383 L384 K385 R386 R387 K388 K389 L393 GLU ALA ALA ALA LEU ALA HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α, β, γ	103.15Å 103.15Å 231.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.34 – 4.00 45.34 – 4.00	Depositor EDS
% Data completeness (in resolution range)	95.6 (45.34-4.00) 95.9 (45.34-4.00)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	0.15	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 4.00Å)	Xtriage
Refinement program	BUSTER-TNT V. 1.1.0	Depositor
R, R_{free}	0.314 , 0.331 0.328 , 0.346	Depositor DCC
R_{free} test set	984 reflections (5.03%)	DCC
Wilson B-factor (Å^2)	106.4	Xtriage
Anisotropy	0.546	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 61.3	EDS
Estimated twinning fraction	0.219 for h,-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Outliers	0 of 19569 reflections	Xtriage
F_o,F_c correlation	0.76	EDS
Total number of atoms	10466	wwPDB-VP
Average B, all atoms (Å^2)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.01% 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

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	0.27	0/2770	0.51	2/3758 (0.1%)
1	B	0.26	0/2775	0.48	2/3765 (0.1%)
2	C	0.25	0/2583	0.59	3/3480 (0.1%)
2	D	0.32	2/2563 (0.1%)	0.57	4/3453 (0.1%)
All	All	0.28	2/10691 (0.0%)	0.54	11/14456 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	B	0	1
2	C	0	24
2	D	0	25
All	All	0	52

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	354	GLU	CD-OE2	6.41	1.32	1.25
2	D	354	GLU	CD-OE1	5.65	1.31	1.25

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	231	ALA	C-N-CD	-7.96	103.09	120.60
2	C	194	TYR	CB-CA-C	-7.67	95.06	110.40
1	A	265	MET	N-CA-C	7.67	131.71	111.00
2	C	234	VAL	N-CA-C	6.29	128.00	111.00
2	D	237	PRO	N-CA-CB	5.93	110.41	103.30

There are no chirality outliers.

5 of 52 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	186	ARG	Peptide
1	A	33	GLY	Peptide
2	C	153	ARG	Peptide
2	C	154	GLY	Peptide
2	C	89	LYS	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	2709	0	2701	298	0
1	B	2714	0	2706	325	0
2	C	2531	0	2509	517	6
2	D	2512	0	2500	495	2
All	All	10466	0	10416	1540	8

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 74.

The worst 5 of 1540 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:B:159:ASN:C	2:D:370:ILE:HD13	1.35	1.42
2:D:214:THR:HG23	2:D:237:PRO:O	1.24	1.37
2:C:315:THR:CG2	2:C:318:GLU:CB	2.02	1.36
2:D:214:THR:CG2	2:D:238:GLU:HA	1.58	1.33
2:D:99:PRO:O	2:D:103:ARG:HG3	1.22	1.32

The worst 5 of 8 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)
2:C:58:ALA:CB	2:C:226:THR:OG1[4_574]	1.62	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:58:ALA:CA	2:C:226:THR:OG1[4_574]	1.75	0.45
2:D:58:ALA:CA	2:D:226:THR:O[3_745]	1.87	0.33
2:D:58:ALA:N	2:D:226:THR:O[3_745]	1.87	0.33
2:C:58:ALA:N	2:C:226:THR:CG2[4_574]	1.90	0.30

5.3 Torsion angles

5.3.1 Protein backbone

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	334/366 (91%)	307 (92%)	25 (8%)	2 (1%)	30 73
1	B	335/366 (92%)	309 (92%)	23 (7%)	3 (1%)	21 66
2	C	311/406 (77%)	262 (84%)	40 (13%)	9 (3%)	6 46
2	D	307/406 (76%)	247 (80%)	52 (17%)	8 (3%)	7 48
All	All	1287/1544 (83%)	1125 (87%)	140 (11%)	22 (2%)	11 56

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	237	PRO
2	D	232	PRO
2	D	237	PRO
2	C	222	THR
2	C	232	PRO

5.3.2 Protein sidechains

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	294/325 (90%)	240 (82%)	54 (18%)	2 15
1	B	294/325 (90%)	243 (83%)	51 (17%)	2 19
2	C	272/362 (75%)	182 (67%)	90 (33%)	0 3
2	D	273/362 (75%)	186 (68%)	87 (32%)	0 3
All	All	1133/1374 (82%)	851 (75%)	282 (25%)	1 7

5 of 282 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	351	ASP
1	B	78	MET
2	D	313	ARG
2	C	360	LEU
2	C	389	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	64	HIS
1	B	202	GLN
2	D	288	ASN
1	B	114	ASN
1	B	228	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

There are no ligands in this entry.

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^{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	A	338/366 (92%)	0.14	8 (2%) 62 51	100, 100, 100, 100	0
1	B	339/366 (92%)	0.01	4 (1%) 81 72	100, 100, 100, 100	0
2	C	317/406 (78%)	0.14	14 (4%) 38 28	20, 100, 100, 100	0
2	D	313/406 (77%)	0.05	5 (1%) 74 64	100, 100, 100, 100	0
All	All	1307/1544 (84%)	0.08	31 (2%) 62 51	20, 100, 100, 100	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	251	SER	6.2
1	A	252	SER	4.3
1	A	56	SER	4.1
2	C	202	ILE	4.0
1	B	85	GLY	3.1

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

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.