



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 11:54 AM GMT

PDB ID : 3Q8W
Title : A b-aminoacyl containing thiazolidine derivative and DPPIV complex
Authors : Lee, J.O.; Song, D.H.
Deposited on : 2011-01-07
Resolution : 3.64 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 ⓘ 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 (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

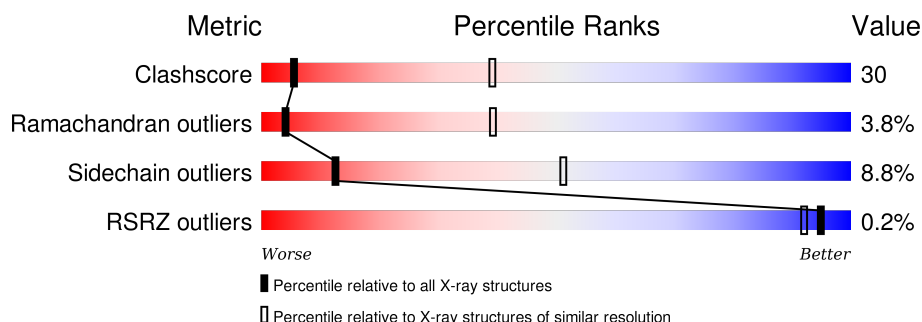
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.64 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))
Clashscore	102246	1130 (3.80-3.48)
Ramachandran outliers	100387	1084 (3.80-3.48)
Sidechain outliers	100360	1083 (3.80-3.48)
RSRZ outliers	91569	1021 (3.80-3.48)

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 $\leq 5\%$. 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	732	<div> <div>45%</div> <div>47%</div> <div>7%</div> </div>
1	B	732	<div> <div>45%</div> <div>47%</div> <div>7%</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
2	AZV	A	1	X	-	-	-
2	AZV	B	1	X	-	-	-

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11972 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.

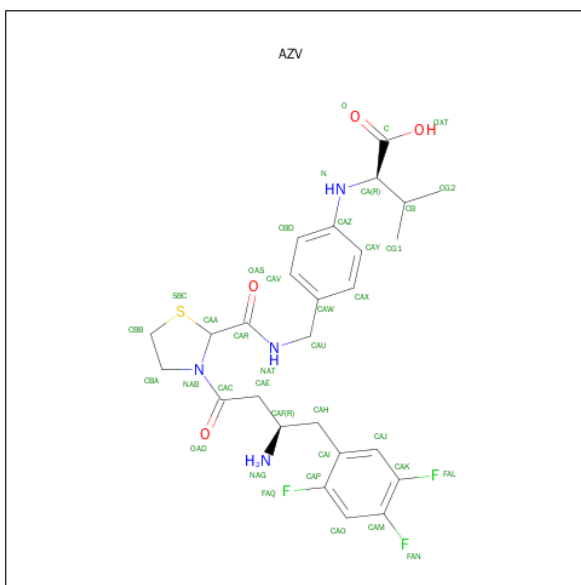
- Molecule 1 is a protein called Dipeptidyl peptidase 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	726	Total	C	N	O	S	0	0	0
			5948	3816	980	1126	26			
1	B	726	Total	C	N	O	S	0	0	0
			5948	3816	980	1126	26			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	765	HIS	-	EXPRESSION TAG	UNP P27487
A	766	HIS	-	EXPRESSION TAG	UNP P27487
A	767	HIS	-	EXPRESSION TAG	UNP P27487
A	768	HIS	-	EXPRESSION TAG	UNP P27487
A	769	HIS	-	EXPRESSION TAG	UNP P27487
A	770	HIS	-	EXPRESSION TAG	UNP P27487
B	765	HIS	-	EXPRESSION TAG	UNP P27487
B	766	HIS	-	EXPRESSION TAG	UNP P27487
B	767	HIS	-	EXPRESSION TAG	UNP P27487
B	768	HIS	-	EXPRESSION TAG	UNP P27487
B	769	HIS	-	EXPRESSION TAG	UNP P27487
B	770	HIS	-	EXPRESSION TAG	UNP P27487

- Molecule 2 is N-(4-{{[(2R)-3-[(3R)-3-AMINO-4-(2,4,5-TRIFLUOROPHENYL)BUTANOYL]-1,3-THIAZOLIDIN-2-YL}CARBONYL)AMINO]METHYL}PHENYL)-D-VALINE (three-letter code: AZV) (formula: C₂₆H₃₁F₃N₄O₄S).

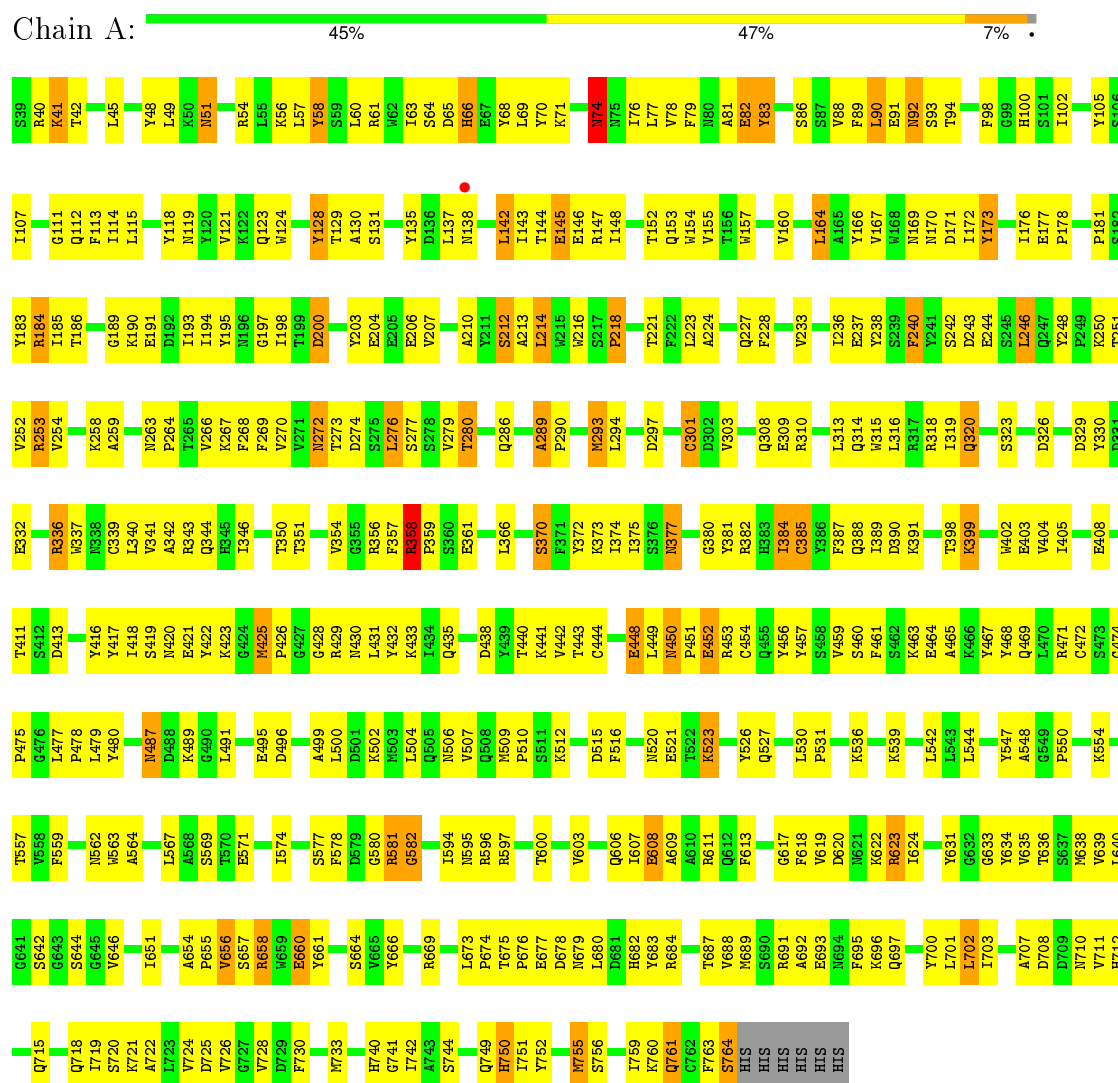


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total 38	C 26	F 3	N 4	O 4	S 1	0	0
2	B	1	Total 38	C 26	F 3	N 4	O 4	S 1	0	0

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: Dipeptidyl peptidase 4



• Molecule 1: Dipeptidyl peptidase 4



L723	L724	D725	V726	G727	V728	W733	D737	H740	G741	I742	A743	S744	T745	T746	A747	H748	Q749	H750	I751	Y752	W755	S756	I759	K760	Q761	G762	F763	S764	HIS	HIS	HIS	HIS	HIS	HIS	S39	R40	K41	T42	L45	Y48	L49	N51	B54	L55	K56	L57	Y58	S59	L60	R61	A62	L63	S64	D65	H66	E67	V68	L69	Y70	K71	R74	H75	L76	L77	V78	F79	N80	A81	E82	Y83	S86	S87	V88	F89	L90	E91	N92	S93	T94	F98	G99	H100	E101	I102	Y105	S106																			
L107	G111	Q112	F113	I114	L115	Y118	N119	K120	V121	K122	Q123	W124	R125	Y126	S127	L128	T129	A130	S131	L137	N138	L142	I143	T144	E145	E146	R147	I148	T152	Q153	W154	V155	T156	W157	V160	L164	A165	Y166	V167	W168	N169	N170	D171	I172	Y173	I176	E177	P178	P181	S182	Y183	R184	I185	T186	G189	K190	E191	D192	I193	L194	Y195	N196	G197	I198	T199	D200	Y203	E204	E205	E206	A210	Y211	S212	A213	W215	W216	S217	P218	T221	F222	L223	A224	Q227	L300	F228	N229	Q308	E309	R310	L313	Q314	W315	L316	R317	Y241	S242	D243	N321	E244	S245	L246	Q247	Y248	P249	K250
T251	V252	R253	V254	K258	A259	N263	T264	T265	V266	K267	F268	F269	V270	T271	N272	T273	D274	S275	L276	S277	Y279	S278	T280	T283	Q286	A289	P290	M293	F292	L294	Y299	L300	C301	Q308	E309	R310	L313	Q314	W315	L316	R317	R318	I319	N321	E244	S245	L246	Q247	Y248	P249	K250																																																						
E332	R336	N338	C339	L340	V341	A342	R343	Q344	H345	T346	T350	T351	V354	G355	R356	F357	R358	P359	S360	E361	L366	S370	F371	V372	K373	I374	S376	R377	G380	Y381	R382	N383	I384	C385	T386	F387	Q388	I389	D390	K391	T398	K399	W402	Y468	Q469	L470	R471	I405	E408																																																								
T411	S412	D413	Y414	L415	Y416	S419	N420	A421	Y422	K423	G424	M425	P426	G427	G428	R429	N430	L431	Y432	F433	I434	Q435	D438	Y439	T440	K441	V442	T443	C444	E448	L449	N450	P451	E452	R453	C454	Q455	Y456	Y457	S458	V459	S460	F461	S462	K463	E464	A465	K466	Y467	A468	Q469	L470	R471	C472	S473																																																		
G474	P475	G476	L477	P478	L479	Y480	S485	V486	N487	D488	K489	E495	D496	A499	L500	D501	K502	N503	L504	Q505	N506	V507	Q508	M509	P510	S511	K512	D515	F516	N520	E521	T522	K523	Y526	R527	P531	K536	K539	L542	L543	L544	Y547	G548	P550	K554																																																												
T557	F558	F559	W563	L567	A568	S569	T570	E571	I574	S577	F578	R581	G582	I594	N595	R596	R597	V603	Q606	I607	E608	A609	A610	R611	Q612	F613	F618	V619	D620	R621	R622	R623	Y631	G632	G633	V634	V635	T636	S637	W638	V639	L640	G641	S642	G643	S644	G645	V646																																																									
I661	A664	P665	V666	S667	R668	W669	E660	S664	V665	V666	R669	L673	P674	T675	P676	E677	Q749	H750	I751	Y752	H681	H682	Y683	R684	T687	V688	R691	F693	A692	E693	N694	F695	K696	Q697	Y700	L701	L702	I703	H704	A707	D708	D709	N710	V711	H712	Q715	Q718	I719	S720	K721	A722																																																						

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	118.41Å 124.50Å 133.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.47 – 3.64 45.46 – 3.64	Depositor EDS
% Data completeness (in resolution range)	88.4 (45.47-3.64) 88.5 (45.46-3.64)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	11.21 (at 3.66Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.282 , 0.324 0.282 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	59.1	Xtriage
Anisotropy	0.901	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 19.9	EDS
Estimated twinning fraction	0.066 for k,h,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.22$	Xtriage
Outliers	0 of 20002 reflections	Xtriage
F_o, F_c correlation	0.81	EDS
Total number of atoms	11972	wwPDB-VP
Average B, all atoms (Å ²)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: AZV

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	0.47	0/6119	0.68	2/8321 (0.0%)
1	B	0.48	0/6119	0.69	2/8321 (0.0%)
All	All	0.48	0/12238	0.69	4/16642 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	319	ILE	N-CA-C	-5.45	96.29	111.00
1	A	319	ILE	N-CA-C	-5.44	96.30	111.00
1	A	240	PHE	N-CA-C	-5.41	96.38	111.00
1	B	240	PHE	N-CA-C	-5.19	97.00	111.00

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	5948	0	5667	357	0
1	B	5948	0	5667	351	0
2	A	38	0	29	1	0
2	B	38	0	29	7	0
All	All	11972	0	11392	692	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 30.

The worst 5 of 692 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:92:ASN:H	1:B:92:ASN:HD22	1.17	0.92
1:A:107:ILE:HG12	1:A:114:ILE:HG12	1.53	0.91
1:A:92:ASN:HD22	1:A:92:ASN:H	1.20	0.89
1:B:107:ILE:HG12	1:B:114:ILE:HG12	1.54	0.89
1:A:310:ARG:HH12	1:A:389:ILE:HD13	1.39	0.86

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	724/732 (99%)	569 (79%)	129 (18%)	26 (4%)	4	41
1	B	724/732 (99%)	571 (79%)	124 (17%)	29 (4%)	4	37
All	All	1448/1464 (99%)	1140 (79%)	253 (18%)	55 (4%)	4	39

5 of 55 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	58	TYR
1	A	181	PRO
1	A	280	THR
1	A	384	ILE
1	A	450	ASN

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	651/657 (99%)	594 (91%)	57 (9%)	12	51
1	B	651/657 (99%)	594 (91%)	57 (9%)	12	51
All	All	1302/1314 (99%)	1188 (91%)	114 (9%)	12	51

5 of 114 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	702	LEU
1	B	126	HIS
1	B	660	GLU
1	A	750	HIS
1	B	41	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	718	GLN
1	B	92	ASN
1	B	712	HIS
1	A	748	HIS
1	B	51	ASN

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

2 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	AZV	A	1	-	36,40,40	2.11	3 (8%)	40,56,56	1.55	5 (12%)
2	AZV	B	1	-	36,40,40	2.15	3 (8%)	40,56,56	1.81	5 (12%)

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	AZV	A	1	-	1/1/8/10	1/29/43/43	0/3/3/3
2	AZV	B	1	-	1/1/8/10	2/29/43/43	0/3/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	AZV	CAA-NAB	-9.34	1.33	1.46
2	A	1	AZV	CAA-NAB	-9.12	1.33	1.46
2	B	1	AZV	CAA-SBC	-7.58	1.75	1.84
2	A	1	AZV	CAA-SBC	-7.32	1.75	1.84
2	B	1	AZV	CAZ-N	-3.34	1.33	1.39

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1	AZV	CAO-CAM-CAK	-4.88	116.70	121.13

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1	AZV	CBA-CBB-SBC	-3.25	98.46	106.84
2	A	1	AZV	CBA-CBB-SBC	-3.19	98.62	106.84
2	A	1	AZV	CAO-CAP-CAI	-2.43	121.03	124.13
2	A	1	AZV	CBA-NAB-CAC	-2.24	119.42	125.58

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	1	AZV	CAA
2	A	1	AZV	CAA

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1	AZV	SBC-CAA-CAR-NAT
2	A	1	AZV	OAS-CAR-CAA-SBC
2	B	1	AZV	OAS-CAR-CAA-SBC

There are no ring outliers.

2 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	AZV	1	0
2	B	1	AZV	7	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(Å ²)	Q<0.9
1	A	726/732 (99%)	-0.17	1 (0%) 95 94	37, 69, 93, 108	0
1	B	726/732 (99%)	-0.10	2 (0%) 94 90	35, 69, 93, 109	0
All	All	1452/1464 (99%)	-0.14	3 (0%) 95 92	35, 69, 93, 109	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	678	ASP	2.2
1	A	138	ASN	2.1
1	B	485	SER	2.1

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, 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(\AA^2)	Q<0.9
2	AZV	A	1	38/38	0.89	0.27	-0.18	47,66,73,74	0
2	AZV	B	1	38/38	0.93	0.23	-0.46	55,62,73,74	0

6.5 Other polymers [i](#)

There are no such residues in this entry.