



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 07:14 AM UTC

PDB ID : 1PVP / pdb_00001pvp
Title : BASIS FOR A SWITCH IN SUBSTRATE SPECIFICITY: CRYSTAL STRUCTURE OF SELECTED VARIANT OF CRE SITE-SPECIFIC RECOMBINASE, ALSHG BOUND TO THE ENGINEERED RECOGNITION SITE LOXM7
Authors : Baldwin, E.P.; Martin, S.S.; Abel, J.; Gelato, K.A.; Kim, H.; Schultz, P.G.; Santoro, S.W.
Deposited on : 2003-06-28
Resolution : 2.35 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

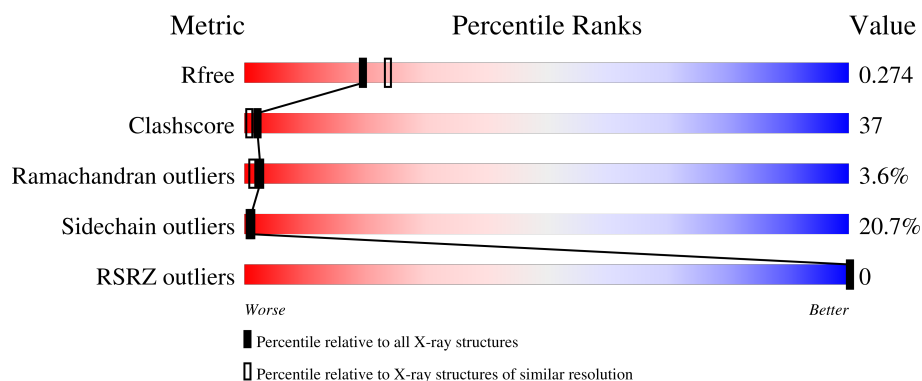
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.35 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}	180053	1596 (2.36-2.36)
Clashscore	190562	1663 (2.36-2.36)
Ramachandran outliers	187476	1646 (2.36-2.36)
Sidechain outliers	187428	1646 (2.36-2.36)
RSRZ outliers	180081	1598 (2.36-2.36)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	C	34	
2	D	34	
3	A	349	
3	B	349	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6738 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 DNA chain called 34-MER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	33	Total	C	N	O	P	0	0	1
			655	317	115	192	31			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	7	DC	DT	engineered mutation	GB M10145
C	8	DT	DC	engineered mutation	GB M10145
C	9	DA	DG	engineered mutation	GB M10145
C	26	DT	DC	engineered mutation	GB M10145
C	27	DA	DG	engineered mutation	GB M10145
C	28	DG	DA	engineered mutation	GB M10145

- Molecule 2 is a DNA chain called 34-MER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	34	Total	C	N	O	P	0	4	0
			774	375	138	224	37			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	7	DC	DT	engineered mutation	GB M10494
D	8	DT	DC	engineered mutation	GB M10494
D	9	DA	DG	engineered mutation	GB M10494
D	26	DT	DC	engineered mutation	GB M10494
D	27	DA	DG	engineered mutation	GB M10494
D	28	DG	DA	engineered mutation	GB M10494

- Molecule 3 is a protein called Recombinase cre.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	323	Total 2546	C 1582	N 486	O 463	S 15	34	0	0
3	B	318	Total 2507	C 1560	N 480	O 452	S 15	27	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	MET	-	initiating methionine	UNP P06956
A	-4	HIS	-	expression tag	UNP P06956
A	-3	HIS	-	expression tag	UNP P06956
A	-2	HIS	-	expression tag	UNP P06956
A	-1	HIS	-	expression tag	UNP P06956
A	0	HIS	-	expression tag	UNP P06956
A	1	HIS	-	expression tag	UNP P06956
A	174	ALA	ILE	engineered mutation	UNP P06956
A	258	LEU	THR	engineered mutation	UNP P06956
A	259	SER	ARG	engineered mutation	UNP P06956
A	262	HIS	GLU	engineered mutation	UNP P06956
A	266	GLY	GLU	engineered mutation	UNP P06956
B	-5	MET	-	initiating methionine	UNP P06956
B	-4	HIS	-	expression tag	UNP P06956
B	-3	HIS	-	expression tag	UNP P06956
B	-2	HIS	-	expression tag	UNP P06956
B	-1	HIS	-	expression tag	UNP P06956
B	0	HIS	-	expression tag	UNP P06956
B	1	HIS	-	expression tag	UNP P06956
B	174	ALA	ILE	engineered mutation	UNP P06956
B	258	LEU	THR	engineered mutation	UNP P06956
B	259	SER	ARG	engineered mutation	UNP P06956
B	262	HIS	GLU	engineered mutation	UNP P06956
B	266	GLY	GLU	engineered mutation	UNP P06956

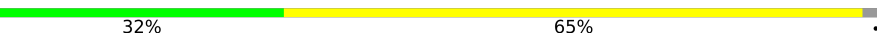
- Molecule 4 is water.

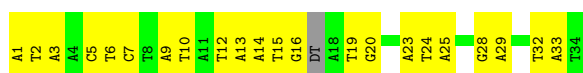
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	31	Total 31	O 31	0	0
4	D	31	Total 31	O 31	0	0
4	A	68	Total 68	O 68	0	0
4	B	126	Total 126	O 126	0	0

3 Residue-property plots

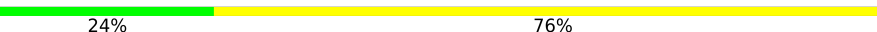
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

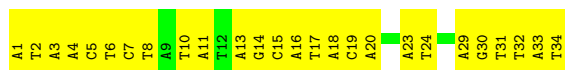
- Molecule 1: 34-MER

Chain C: 

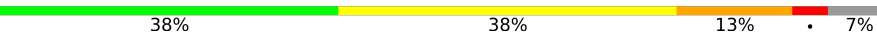


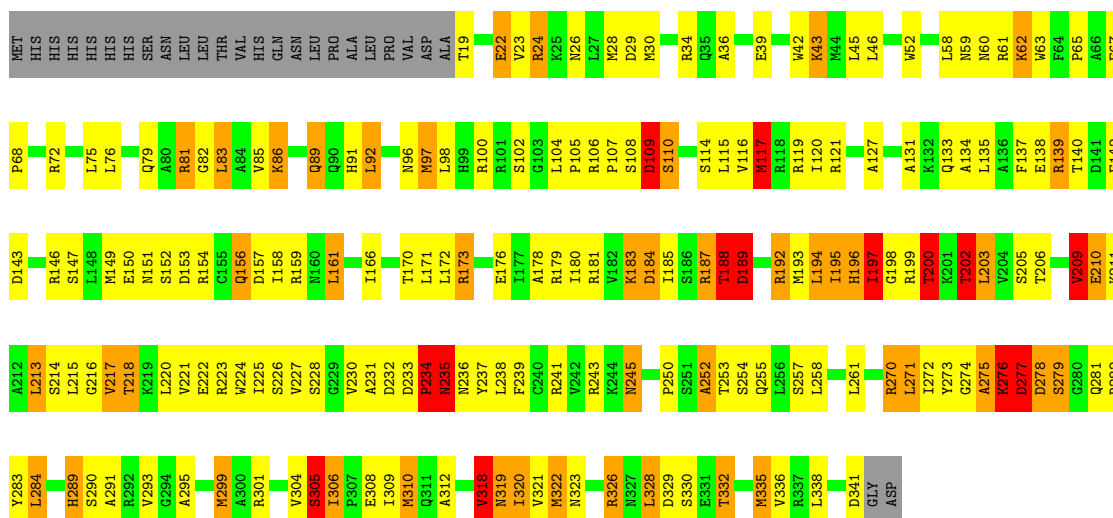
- Molecule 2: 34-MER

Chain D: 

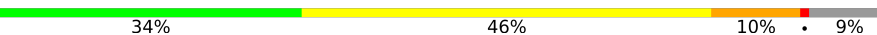


- Molecule 3: Recombinase cre

Chain A: 



- Molecule 3: Recombinase cre

Chain B: 

W286	R292	V217	Q144	E69	MET
S287	V293	T218	V145		HIS
G288	G294	K219	R146	R72	HIS
H289	A295	L220		D73	HIS
	A296	V221	E150	Y74	HIS
	R297	E222			HIS
	D298	R223	R154	Y77	SER
	M299	W224	C155		ASN
A300	A299		Q156	R81	LEU
R301	R301	V230	D157		LEU
		A231	I158	A84	THR
		D232	R159	V85	VAL
		D233	H160	K86	HIS
		P234	L161	T87	GLN
		N235		I88	ASN
		N236	I166	Q89	LEU
		Y237	A167		PRO
				L92	ALA
					ALA
		R241	T170		LEU
		V242	L171	N96	PRO
		R243	L172		VAL
		K244	R173	H99	ASP
		N245	A174	R100	ASP
		G246	A175	R101	ALA
		V247	E176	S102	T19
G313		A248	I177	G103	S20
G314		A249	A178	L104	D21
W315		P250	R179	P105	E22
T316		S251	I180	R106	
N317		A252	R181	P107	K25
V318		T253	V182	S108	W26
N319		S254	K183	D109	L27
I320		Q255	D184	S110	M28
V321		L256	I185		D29
N322		S257	S186	V113	M30
N323		L258	R187		F31
Y324		I325	T188	V116	R32
I325		A260		M117	D33
		L261	M193	R118	R34
R326			L194	R119	
ASN			I195	I120	H40
LEU			H196	R121	T41
ASP			I197	K122	W42
GLU			G198	E123	K43
T332		T268	R199	H124	M44
G333		H269	T200	V125	L45
A334		R270	K201	D126	L46
M335		L271	T202	A127	S47
V336		I272	T203	G128	V48
		Y273	L263	E129	C49
R337		G274	V204	R130	R50
L338		A275	S205	A131	
L339		K276	T206	K132	N60
E340		D277	A207	Q133	R61
D341		D278	G208	K134	K62
GLY		S279	V209	A134	W63
ASP		G280	E210	E138	F64
		Q281	K211	R139	P65
		R282	A212	T140	
		Y283	L213		P68
		L284			
		A285			

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	107.78Å 121.47Å 180.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	5.00 – 2.35 5.00 – 2.35	Depositor EDS
% Data completeness (in resolution range)	95.0 (5.00-2.35) 84.6 (5.00-2.35)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.78 (at 2.34Å)	Xtriage
Refinement program	TNT	Depositor
R, R_{free}	0.232 , 0.294 0.235 , 0.274	Depositor DCC
R_{free} test set	2180 reflections (4.65%)	wwPDB-VP
Wilson B-factor (Å ²)	47.6	Xtriage
Anisotropy	0.301	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 112.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6738	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	C	0.25	0/733	1.14	1/1129 (0.1%)
2	D	0.34	0/868	1.01	0/1337
3	A	0.50	0/2588	1.14	22/3490 (0.6%)
3	B	0.54	0/2548	1.15	19/3434 (0.6%)
All	All	0.47	0/6737	1.13	42/9390 (0.4%)

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
3	A	1	0

There are no bond length outliers.

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	322	MET	N-CA-C	11.71	124.05	111.28
3	A	202	THR	N-CA-C	9.96	132.01	110.80
3	A	235	ASN	N-CA-C	-9.88	100.91	113.16
3	A	326	ARG	N-CA-C	9.57	122.92	111.82
3	B	188	THR	N-CA-C	-8.98	97.56	110.59

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	202	THR	CA

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	C	655	0	368	28	0
2	D	774	0	434	62	0
3	A	2546	0	2566	205	0
3	B	2507	0	2533	177	0
4	A	68	0	0	13	0
4	B	126	0	0	6	0
4	C	31	0	0	1	0
4	D	31	0	0	3	0
All	All	6738	0	5901	448	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 37.

The worst 5 of 448 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)
2:D:33:DA:H2''	2:D:34:DT:H5''	1.30	1.11
2:D:15:DC:H2'	2:D:16[B]:DA:C8	1.88	1.07
3:B:306:ILE:HD13	3:B:322:MET:HE1	1.32	1.06
2:D:15:DC:C2'	2:D:16[B]:DA:C8	2.47	0.97
3:B:193:MET:HE2	3:B:213:LEU:HD12	1.43	0.96

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	
3	A	321/349 (92%)	269 (84%)	35 (11%)	17 (5%)	1	0
3	B	314/349 (90%)	290 (92%)	18 (6%)	6 (2%)	6	5
All	All	635/698 (91%)	559 (88%)	53 (8%)	23 (4%)	2	1

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	202	THR
3	A	210	GLU
3	A	234	PRO
3	A	275	ALA
3	A	278	ASP

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	
3	A	268/291 (92%)	204 (76%)	64 (24%)	1	0
3	B	263/291 (90%)	217 (82%)	46 (18%)	2	1
All	All	531/582 (91%)	421 (79%)	110 (21%)	1	1

5 of 110 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	323	ASN
3	B	46	LEU
3	B	340	GLU
3	B	305	SER
3	A	328	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	156	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	B	289	HIS
3	B	319	ASN
3	B	317	ASN
3	A	245	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	D	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	15:DC	O3'	16[B]:DA	P	1.84
1	D	19[B]:DC	O3'	20:DA	P	1.26

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	C	33/34 (97%)	-1.01	0 100 100	32, 54, 85, 100	0
2	D	34/34 (100%)	-1.08	0 100 100	27, 47, 59, 78	4 (11%)
3	A	322/349 (92%)	-0.41	0 100 100	27, 59, 96, 99	9 (2%)
3	B	318/349 (91%)	-0.74	0 100 100	25, 45, 83, 98	7 (2%)
All	All	707/766 (92%)	-0.62	0 100 100	25, 53, 90, 100	20 (2%)

There are no RSRZ outliers to report.

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.