



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 06:41 AM GMT

PDB ID : 3D7U  
Title : Structural basis for the recognition of c-Src by its inactivator Csk  
Authors : Levinson, N.M.; Seeliger, M.A.; Cole, P.A.; Kuriyan, J.  
Deposited on : 2008-05-21  
Resolution : 4.11 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

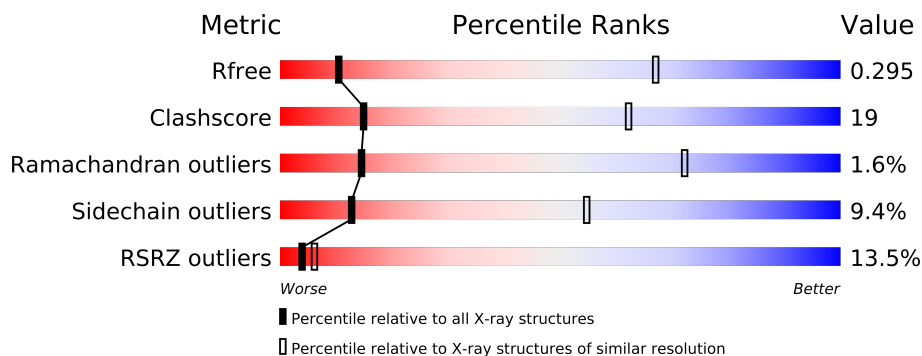
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 4.11 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))
$R_{free}$	66092	1003 (4.76-3.48)
Clashscore	79885	1248 (4.70-3.50)
Ramachandran outliers	78287	1183 (4.70-3.50)
Sidechain outliers	78261	1168 (4.70-3.50)
RSRZ outliers	66119	1003 (4.76-3.48)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	263	
1	C	263	
2	B	277	
2	D	277	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8153 atoms, of which 0 are hydrogen and 0 are deuterium.

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 CSK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	257	Total	C	N	O	S	0	0	0
			2038	1306	349	370	13			
1	C	257	Total	C	N	O	S	0	0	0
			2038	1306	349	370	13			

- Molecule 2 is a protein called Proto-oncogene tyrosine-protein kinase Src.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	258	Total	C	N	O	S	0	0	0
			2080	1334	347	382	17			
2	D	248	Total	C	N	O	S	0	0	0
			1997	1283	334	363	17			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	257	LYS	-	EXPRESSION TAG	UNP P00523
B	258	ASP	-	EXPRESSION TAG	UNP P00523
B	259	ALA	-	EXPRESSION TAG	UNP P00523
B	524	GLU	-	EXPRESSION TAG	UNP P00523
B	525	PRO	-	EXPRESSION TAG	UNP P00523
B	526	GLN	-	EXPRESSION TAG	UNP P00523
B	527	TYR	-	EXPRESSION TAG	UNP P00523
B	528	GLN	-	EXPRESSION TAG	UNP P00523
B	529	PRO	-	EXPRESSION TAG	UNP P00523
B	530	GLY	-	EXPRESSION TAG	UNP P00523
B	531	GLU	-	EXPRESSION TAG	UNP P00523
B	532	ASN	-	EXPRESSION TAG	UNP P00523
B	533	LEU	-	EXPRESSION TAG	UNP P00523
D	257	LYS	-	EXPRESSION TAG	UNP P00523
D	258	ASP	-	EXPRESSION TAG	UNP P00523
D	259	ALA	-	EXPRESSION TAG	UNP P00523

*Continued on next page...*

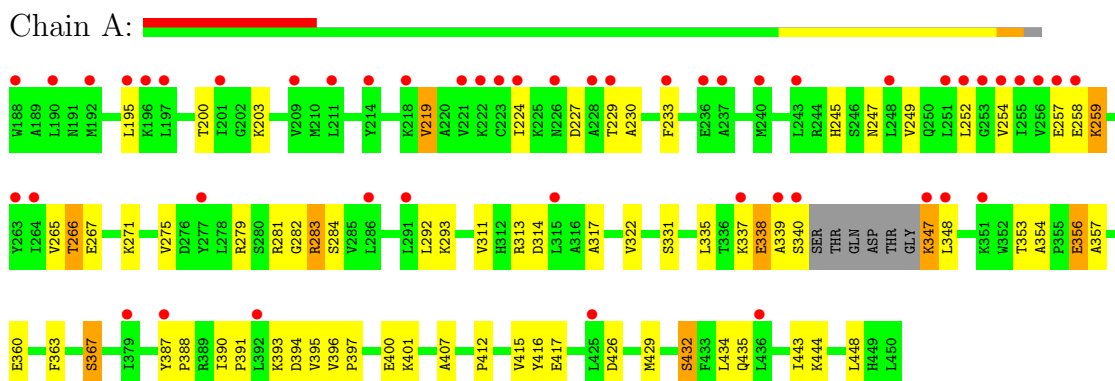
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	524	GLU	-	EXPRESSION TAG	UNP P00523
D	525	PRO	-	EXPRESSION TAG	UNP P00523
D	526	GLN	-	EXPRESSION TAG	UNP P00523
D	527	TYR	-	EXPRESSION TAG	UNP P00523
D	528	GLN	-	EXPRESSION TAG	UNP P00523
D	529	PRO	-	EXPRESSION TAG	UNP P00523
D	530	GLY	-	EXPRESSION TAG	UNP P00523
D	531	GLU	-	EXPRESSION TAG	UNP P00523
D	532	ASN	-	EXPRESSION TAG	UNP P00523
D	533	LEU	-	EXPRESSION TAG	UNP P00523

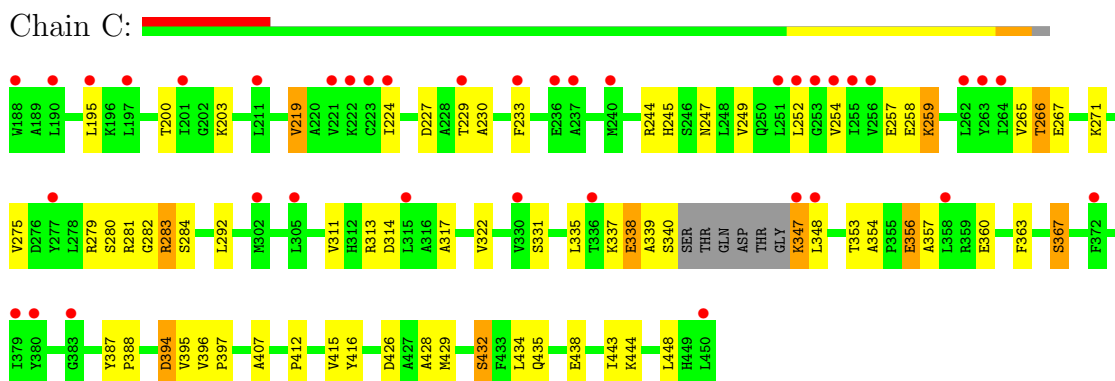
### 3 Residue-property plots

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.

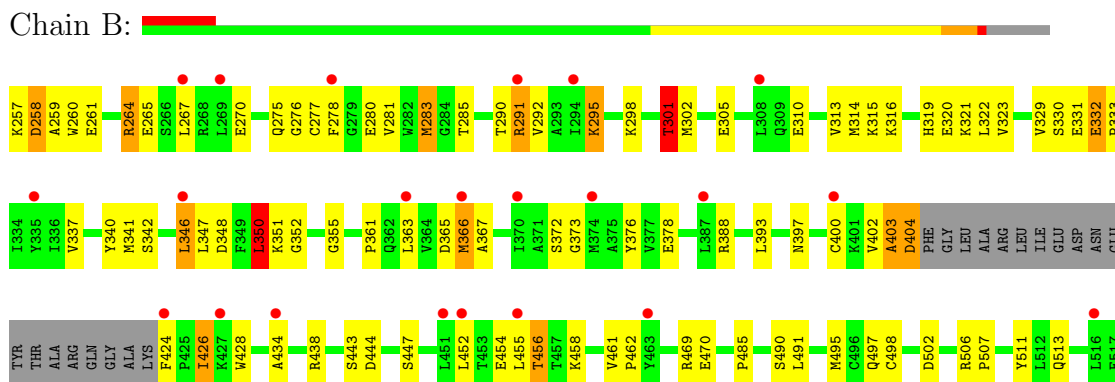
#### • Molecule 1: Tyrosine-protein kinase CSK

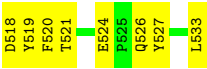


#### • Molecule 1: Tyrosine-protein kinase CSK

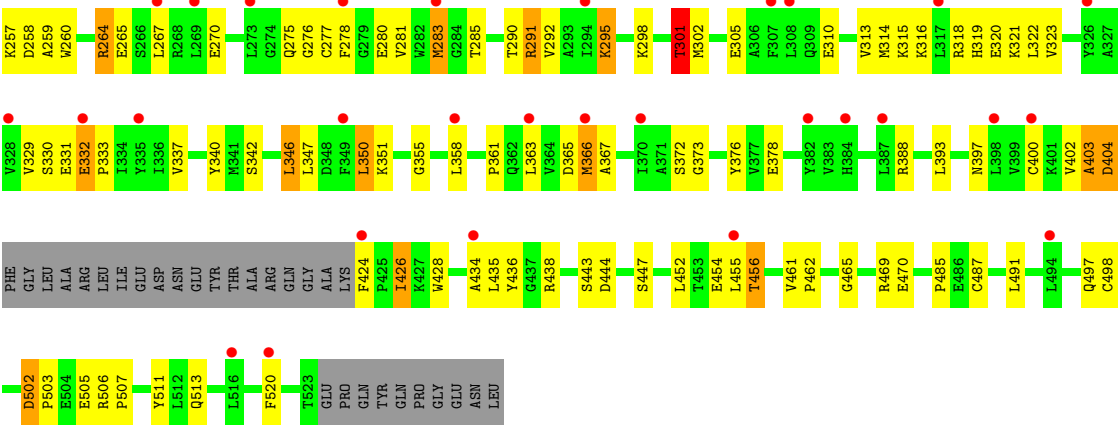


#### • Molecule 2: Proto-oncogene tyrosine-protein kinase Src





● Molecule 2: Proto-oncogene tyrosine-protein kinase Src



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	135.84Å 135.84Å 129.51Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.87 – 4.11 46.87 – 4.11	Depositor EDS
% Data completeness (in resolution range)	98.5 (46.87-4.11) 98.6 (46.87-4.11)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.01 (at 4.14Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.294 , 0.302 0.289 , 0.295	Depositor DCC
$R_{free}$ test set	1053 reflections (5.14%)	DCC
Wilson B-factor (Å <sup>2</sup> )	162.2	Xtriage
Anisotropy	0.301	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 129.5	EDS
Estimated twinning fraction	0.034 for -h,-k,l 0.078 for h,-h-k,-l 0.036 for -k,-h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 20508 reflections	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	8153	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	191.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	0.45	1/2080 (0.0%)	0.64	0/2806
1	C	0.45	1/2080 (0.0%)	0.64	0/2806
2	B	0.76	2/2130 (0.1%)	0.89	8/2882 (0.3%)
2	D	0.82	2/2044 (0.1%)	0.89	6/2765 (0.2%)
All	All	0.64	6/8334 (0.1%)	0.77	14/11259 (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
2	B	0	1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	350	LEU	C-N	15.92	1.70	1.34
2	B	350	LEU	C-N	8.27	1.53	1.34
1	A	432	SER	C-O	5.77	1.34	1.23
1	C	432	SER	C-O	5.76	1.34	1.23
2	D	342	SER	C-O	5.71	1.34	1.23

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	350	LEU	CA-C-N	-8.24	99.08	117.20
2	B	350	LEU	CA-C-N	-8.08	99.43	117.20
2	D	365	ASP	CB-CG-OD2	6.89	124.50	118.30
2	B	365	ASP	CB-CG-OD2	6.88	124.49	118.30
2	D	258	ASP	CB-CG-OD2	6.84	124.45	118.30

There are no chirality outliers.



All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	350	LEU	Mainchain

## 5.2 Close contacts ⓘ

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2038	0	2064	92	9
1	C	2038	0	2064	76	12
2	B	2080	0	2065	103	1
2	D	1997	0	1993	108	4
All	All	8153	0	8186	312	13

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 19.

The worst 5 of 312 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:D:350:LEU:C	2:D:351:LYS:N	1.70	1.44
1:C:279:ARG:HD3	2:D:511:TYR:CE2	1.54	1.41
1:A:279:ARG:HD3	2:B:511:TYR:CE2	1.67	1.29
1:A:400:GLU:OE1	2:D:465:GLY:CA	1.79	1.27
1:C:279:ARG:CD	2:D:511:TYR:CE2	2.18	1.26

The worst 5 of 13 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	Distance(Å)	Clash(Å)
1:A:401:LYS:O	1:C:438:GLU:CD[3_664]	0.35	1.85
1:A:401:LYS:O	1:C:438:GLU:OE2[3_664]	1.08	1.12
1:A:401:LYS:O	1:C:438:GLU:OE1[3_664]	1.30	0.90
1:A:401:LYS:C	1:C:438:GLU:OE1[3_664]	1.52	0.68
1:A:401:LYS:C	1:C:438:GLU:CD[3_664]	1.54	0.66

## 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	253/263 (96%)	237 (94%)	13 (5%)	3 (1%)	19	77
1	C	253/263 (96%)	237 (94%)	13 (5%)	3 (1%)	19	77
2	B	252/277 (91%)	219 (87%)	28 (11%)	5 (2%)	11	68
2	D	242/277 (87%)	209 (86%)	28 (12%)	5 (2%)	11	66
All	All	1000/1080 (93%)	902 (90%)	82 (8%)	16 (2%)	14	72

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	339	ALA
1	A	395	VAL
2	B	403	ALA
1	C	339	ALA
1	C	395	VAL

### 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	220/225 (98%)	210 (96%)	10 (4%)	38	82
1	C	220/225 (98%)	209 (95%)	11 (5%)	34	80
2	B	225/239 (94%)	194 (86%)	31 (14%)	5	34
2	D	216/239 (90%)	185 (86%)	31 (14%)	5	32
All	All	881/928 (95%)	798 (91%)	83 (9%)	13	55

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	470	GLU
1	C	283	ARG
2	D	447	SER
2	B	491	LEU
1	C	219	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	526	GLN
2	B	532	ASN
1	C	442	HIS
2	B	474	GLN
1	C	446	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	257/263 (97%)	1.29	49 (19%) <b>2</b> <b>3</b>	143, 178, 256, 262	0
1	C	257/263 (97%)	1.07	38 (14%) <b>3</b> <b>5</b>	149, 189, 243, 252	0
2	B	258/277 (93%)	0.82	22 (8%) <b>11</b> <b>13</b>	140, 182, 212, 232	0
2	D	248/277 (89%)	0.91	29 (11%) <b>5</b> <b>8</b>	143, 198, 221, 229	0
All	All	1020/1080 (94%)	1.02	138 (13%) <b>4</b> <b>6</b>	140, 186, 244, 262	0

The worst 5 of 138 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	224	ILE	7.8
1	A	222	LYS	6.8
1	A	197	LEU	6.4
1	A	195	LEU	6.3
1	C	223	CYS	6.2

### 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 ⓘ

There are no such residues in this entry.