Integrative Structure Validation Report

October 09, 2025 - 04:37 PM PDT

The following software was used in the production of this report:

IHMValidation Version 3.0
Python-IHM Version 2.5
MolProbity Version 4.5.2
ATSAS Version 3.2.1 (r14885)

PDB ID	8ZZR | pdb_00008zzr
PDB-Dev ID	PDBDEV_00000027
Structure Title	A metastable contact and structural disorder in the estrogen receptor transactivation domain
Structure Authors	Peng Y; Cao S; Kiselar J; Xiao X; Du Z; Hsien A; Ko S; Chen Y; Agrawal P; Zheng W; Shi W; Jiang W; Yang L; Chance MR; Surewicz WK; Buck M; Yang S
Deposited on	2018-09-26

This is a PDB-IHM Structure Validation Report.

We welcome your comments at helpdesk@pdb-ihm.org

A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.

List of references used to build this report is available here.

1. Overview

1.1. Summary

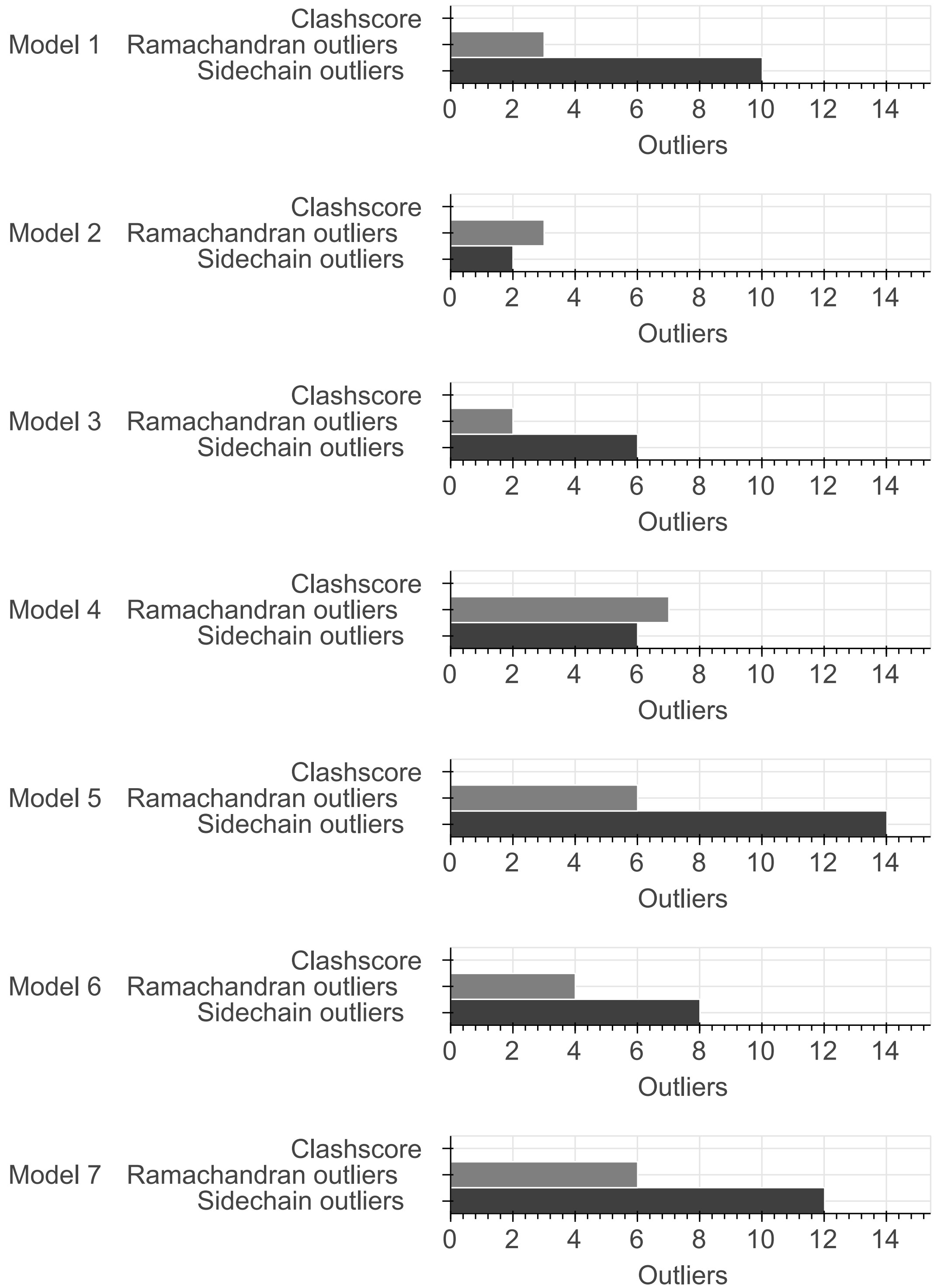
This entry consists of 10 model(s). A total of 2 dataset(s) were used to build this entry.

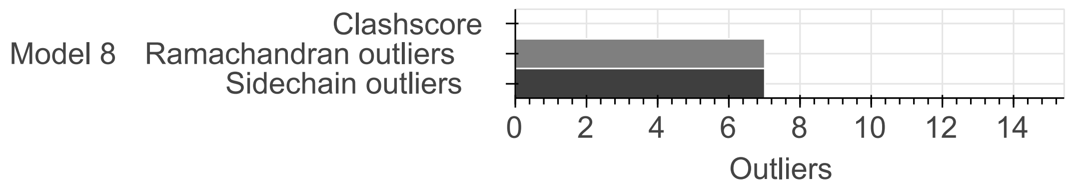
Name	Type	Count
Hydroxyl radical footprinting data	Experimental data	1
SAS data	Experimental data	1

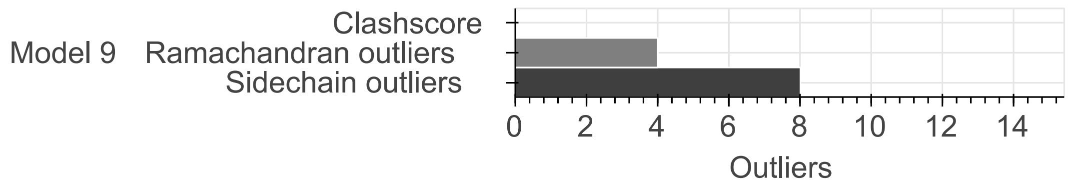
1.2. Overall quality

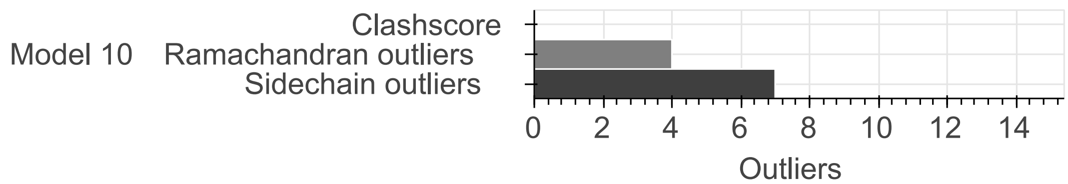
This validation report contains model quality assessments for all structures, data quality and fit to model assessments for SAS and crosslinking-MS datasets. Data quality and fit to model assessments for other datasets and model uncertainty are under development. Number of plots is limited to 256.

Model Quality: MolProbity Analysis

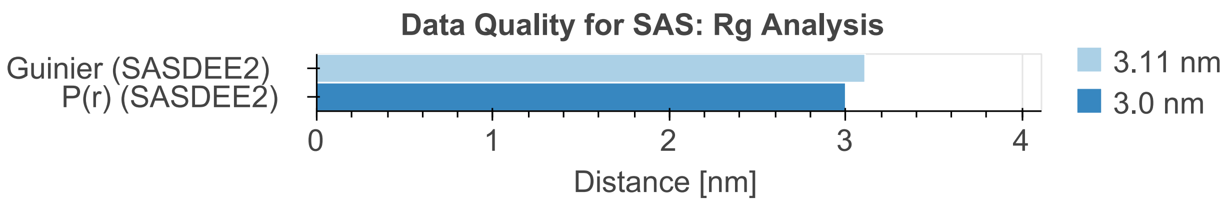




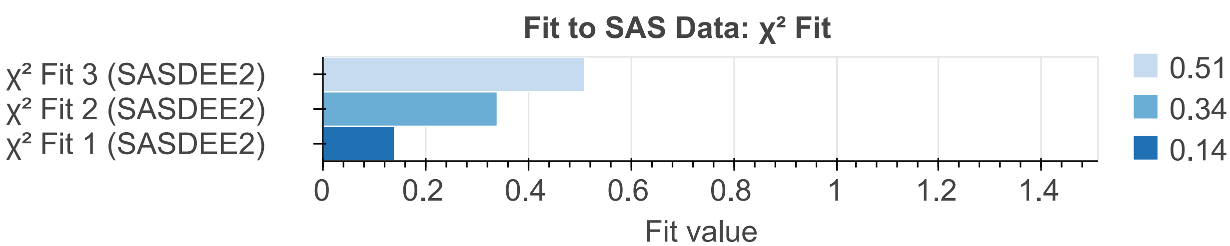




Data Quality



Fit to Data Used for Modeling



2. Model Details

2.1. Ensemble information

This entry consists of 0 distinct ensemble(s).

2.2. Representation

This entry has 1 representation(s).

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
1	1-10	1	Estrogen receptor	A	184	1-184	-	100.00 / 0.00	Atomic

2.3. Datasets used for modeling

There are 2 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	SAS data	SASBDB	SASDEE2
2	Hydroxyl radical footprinting data	Not available	10.1016/j.str.2018.10.026

2.4. Methodology and software

This entry is a result of 1 distinct protocol(s).

Step number	Protocol ID	Method name	Method type	Method description	Number of computed models	Multi state modeling	Multi scale modeling
1	1	Modeling estrogen receptor N-terminal domain	Not available	Not available	Not available	False	False

There is 1 software package reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	iSPOT	Not available	model building	http://www.theyanglab.org/ispot/index.html

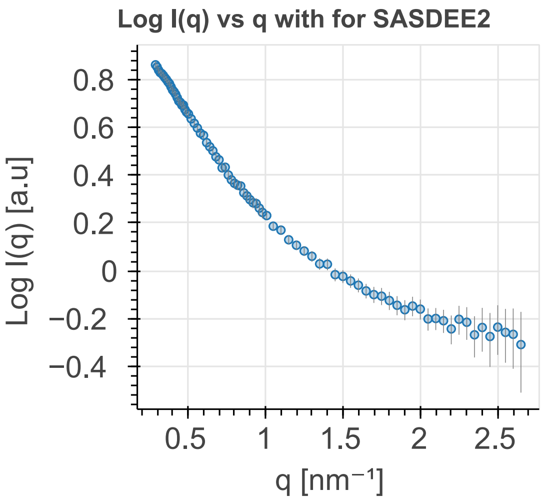
3. Data quality

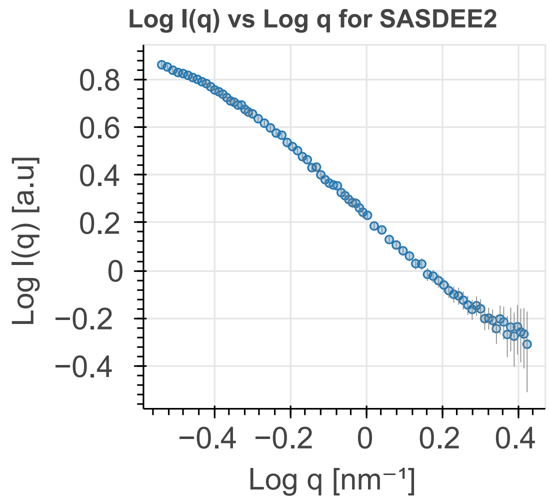
3.1. SAS

3.1.1. Scattering profile

SAS data used in this integrative model was obtained from 1 deposited SASBDB entry (entries).

Scattering profile for SASDEE2: data from solutions of biological macromolecules are presented as both log I(q) vs q and log I(q) vs log (q) based on SAS validation task force (SASvtf) recommendations. I(q) is the intensity (in arbitrary units) and q is the modulus of the scattering vector.





3.1.2. Key experimental estimates

Molecular weight (MW) estimates from experiments and analysis: Theoretical MW can be compared to SAS-derived values using the forward scatter (*I(0)*) and the known concentration and partial specific volume of the scattering particle, or as estimated from the Porod volume and partial specific volume (Trewhella et al., 2017, Trewhella et al., 2023).

SASDB ID	Chemical composition MW	Standard MW	Porod Volume/MW
SASDEE2	20.2 kDa	Not available	Not available

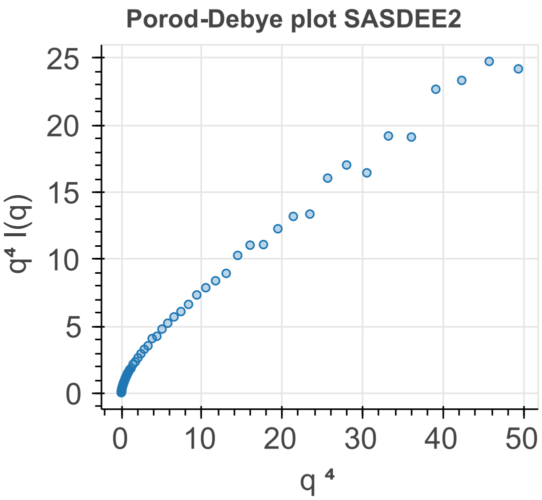
Volume estimates from experiments and analysis: estimated volume can be compared to Porod volume obtained from scattering profiles.

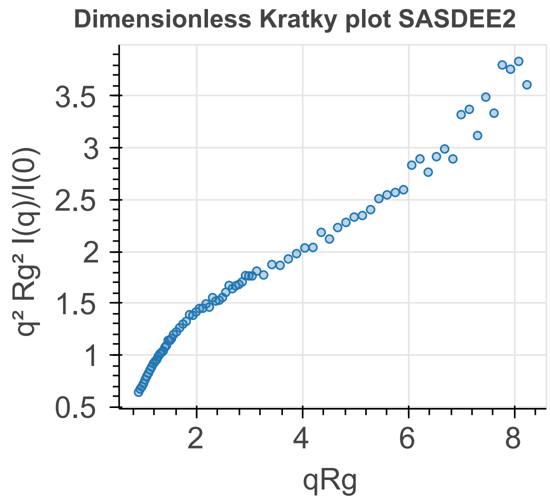
SASDB ID	Estimated Volume	Porod Volume	Specific Volume	Sample Contrast	Sample Concentration
SASDEE2	Not available	Not available	Not available	Not available	2.50 mg/mL

3.1.3. Flexibility analysis

In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas flexible-modular, fully unfolded domains or extended/stiff rodshaped domains lack a discernible plateau (Rambo and Tainer 2013). A bell-shaped Kratky plot ($q^2I(q)$ *vs.* q) with a well-defined maximum is observed for compact/folded structures. For partially flexible/modular or extended structures the Kratky plot can show multiple maxima and/or an increase in intensity at higher q-values depending on the degree of flexibility and extension. Fully intrinsically disordered structures yield a Kratky plot that systematically increases with increasing q values and will be near linear for highly extended molecules. The dimensionless Kratky plot ($(qR_g)^2I(q)$ *vs.* qR_g) is useful for quantifying differences in shape and foldedness among scattering objects of different sizes (Trewhella et. al., 2023).

Flexibility analysis for SASDEE2



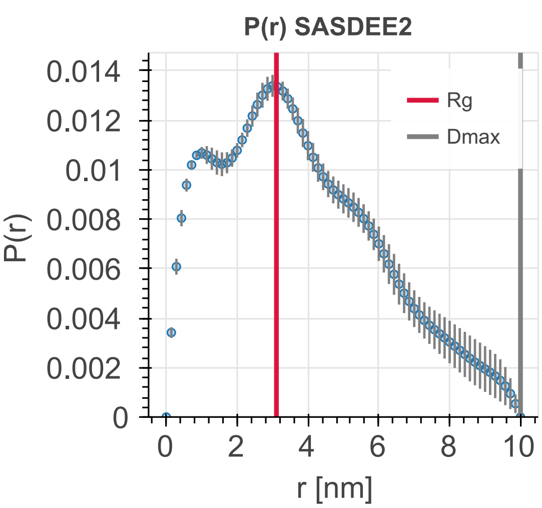


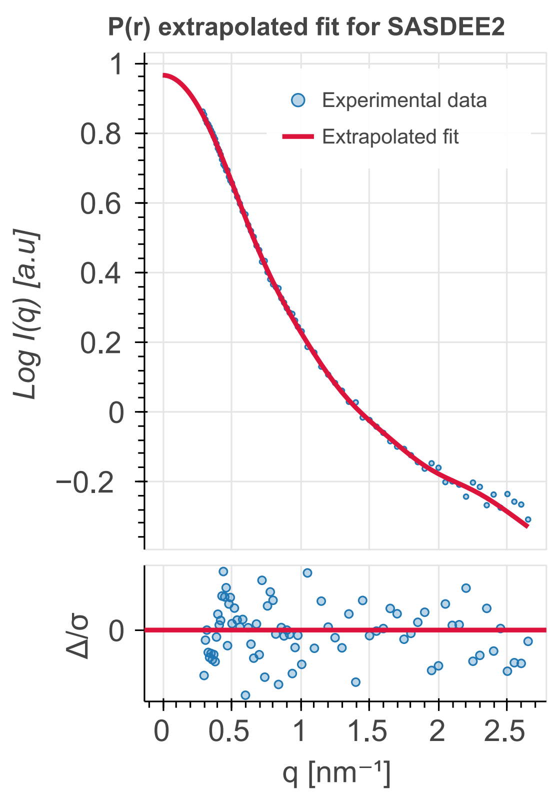
3.1.4. Pair-distance distribution analysis

The the atom-pair distance distribution function (PDDF) or *P(r)* represents the distribution of distances between all pairs of atoms within the particle weighted by the respective scattering contrasts (Moore, 1980). The second moment of *P(r)* yields the radius of gyration (Rg), which is a measure of the overall size and shape of a macromolecule (i.e. the spatial distribution of volume elements). A protein with a smaller R_g is more compact than a protein with a larger R_g, provided both have the same molecular weight.

SASDB ID	Software used	D_{max}	D_{max} error	R_g	R_g error
SASDEE2	GNOM 4.6	10.000 nm	Not available	3.110 nm	Not available

P(r) for SASDEE2: The value of P(r) should be zero beyond $r=D_{max}$.

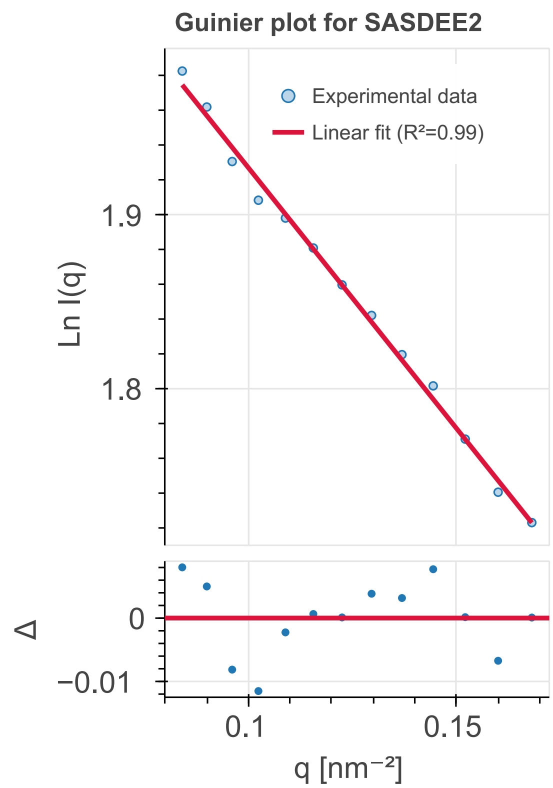




3.1.5. Guinier analysis

The linearity of the Guinier plot (ln(q) *vs.* q^2) at very-low angle ($qRg < 1.3$) is a sensitive indicator of the quality of the sample in relation to its homogeneity; a linear Guinier plot is a necessary but not sufficient demonstration that a solution contains monodisperse particles of the same size. Deviations from linearity can point to strong interference effects from particle attraction or repulsion, polydispersity of the samples, or improper background subtraction (Feigin et al., 2013). Residual difference plots and Pearson correlation coefficient determination (R^2) are measures to assess quality of the linear fit to the Guinier region. A perfect fit has an R^2 value of 1. Residual values should be equally and randomly spaced around the horizontal axis with no evident systematic upward or downward curvature.

SASDB ID	R_g	R_g error	MW	MW error
SASDEE2	3.00 nm	0.16 nm	Not available	Not available



3.4. Hydroxyl radical footprinting

Validation for this section is under development.

4. Model quality

For models with atomic structures, MolProbity analysis is performed. For models with coarse-grained or multi-scale structures, excluded volume analysis is performed.

4.1b. MolProbity Analysis

Excluded volume satisfaction for the models in the entry are listed below. The Analysed column shows the number of particle-partice or particle-atom pairs for which excluded volume was analysed.

Standard geometry: bond outliers

There are 247 bond length outliers in this entry (1.71% of 14440 assessed bonds). A summary is provided below. The output is limited to 100 rows.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	142	ARG	CZ-NH2	10.91	1.19	1.33	3	4
A	112	HIS	CG-ND1	9.68	1.48	1.38	8	3
A	112	HIS	CE1-NE2	7.47	1.40	1.32	7	2
A	112	HIS	ND1-CE1	7.39	1.39	1.32	5	7
A	147	PRO	CA-CB	7.27	1.68	1.53	10	1
A	109	MET	C-N	7.20	1.43	1.33	10	2
A	162	ARG	CD-NE	7.01	1.56	1.46	3	1
A	129	PRO	N-CD	6.72	1.38	1.47	1	1
A	16	HIS	CE1-NE2	6.61	1.39	1.32	7	4
A	89	PHE	C-N	6.41	1.42	1.33	6	1
A	124	HIS	ND1-CE1	6.37	1.38	1.32	5	1
A	28	ARG	NE-CZ	6.27	1.39	1.33	7	2
A	16	HIS	ND1-CE1	6.22	1.38	1.32	4	6
A	158	ARG	NE-CZ	6.17	1.39	1.33	10	3
A	37	ARG	CZ-NH2	6.17	1.25	1.33	9	1
A	175	ALA	C-N	6.12	1.41	1.33	9	2
A	124	HIS	CG-ND1	6.06	1.44	1.38	10	3
A	38	PRO	N-CA	6.04	1.38	1.47	2	1
A	34	PRO	C-N	6.02	1.41	1.33	1	1
A	127	GLN	C-N	5.98	1.41	1.33	5	1
A	59	ALA	CA-C	5.92	1.65	1.52	9	1
A	124	HIS	CE1-NE2	5.91	1.38	1.32	8	3
A	151	ARG	CZ-NH2	5.89	1.25	1.33	6	2
A	98	PRO	N-CA	5.84	1.38	1.47	2	1
A	148	ALA	C-N	5.79	1.41	1.33	10	1
A	119	PRO	N-CA	5.76	1.38	1.47	6	2
A	115	PRO	N-CD	5.67	1.39	1.47	5	2
A	2	THR	CB-OG1	5.62	1.34	1.43	8	2
A	84	SER	CB-OG	5.59	1.53	1.42	9	1
A	96	GLY	N-CA	5.59	1.54	1.45	10	2
A	146	PRO	N-CD	5.57	1.40	1.47	2	2
A	141	VAL	C-N	5.51	1.41	1.33	9	1
A	51	VAL	C-N	5.51	1.41	1.33	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	29	PRO	N-CA	5.50	1.55	1.47	3	1
A	89	PHE	CA-CB	5.49	1.64	1.53	9	1
A	79	PRO	N-CA	5.47	1.55	1.47	5	3
A	135	GLU	CA-C	5.47	1.64	1.52	1	1
A	7	THR	CB-OG1	5.44	1.35	1.43	7	3
A	6	HIS	CG-CD2	5.44	1.41	1.35	7	3
A	6	HIS	CG-ND1	5.44	1.44	1.38	1	2
A	74	GLY	C-N	5.35	1.40	1.33	10	1
A	36	GLU	C-N	5.26	1.40	1.33	2	2
A	182	THR	CB-OG1	5.22	1.35	1.43	2	1
A	66	ALA	C-N	5.21	1.40	1.33	2	2
A	144	ALA	CA-C	5.18	1.63	1.52	7	1
A	123	PRO	N-CA	5.17	1.39	1.47	6	1
A	105	PRO	N-CD	5.13	1.54	1.47	6	3
A	76	THR	CB-OG1	5.12	1.35	1.43	1	1
A	158	ARG	CZ-NH2	5.12	1.26	1.33	7	1
A	136	PRO	N-CD	5.10	1.40	1.47	6	1
A	37	ARG	NE-CZ	5.09	1.38	1.33	4	1
A	17	GLN	N-CA	5.09	1.55	1.46	6	1
A	37	ARG	CD-NE	5.07	1.53	1.46	8	1
A	28	ARG	CZ-NH2	5.06	1.26	1.33	3	2
A	98	PRO	C-N	5.05	1.42	1.34	4	1
A	157	ARG	NE-CZ	5.05	1.38	1.33	4	1
A	28	ARG	CZ-NH1	5.05	1.25	1.32	6	1
A	154	SER	C-N	5.04	1.40	1.33	6	2
A	60	TYR	CA-C	5.00	1.63	1.52	2	1
A	9	ALA	C-N	4.99	1.26	1.33	7	1
A	34	PRO	N-CD	4.99	1.40	1.47	7	2
A	178	SER	N-CA	4.98	1.55	1.46	2	1
A	107	PRO	N-CD	4.97	1.40	1.47	2	1
A	68	ALA	C-N	4.97	1.26	1.33	3	1
A	20	GLY	CA-C	4.96	1.43	1.52	8	2
A	23	LEU	CA-C	4.93	1.63	1.52	7	1
A	140	THR	C-N	4.93	1.40	1.33	2	1
A	6	HIS	CE1-NE2	4.89	1.37	1.32	7	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	104	SER	CB-OG	4.88	1.32	1.42	1	1
A	4	THR	CB-OG1	4.86	1.36	1.43	5	1
A	151	ARG	NE-CZ	4.86	1.38	1.33	9	1
A	45	ASP	C-N	4.85	1.40	1.33	4	1
A	133	GLU	C-N	4.83	1.40	1.33	10	1
A	2	THR	C-N	4.82	1.26	1.33	5	1
A	162	ARG	CZ-NH2	4.80	1.27	1.33	10	2
A	105	PRO	N-CA	4.80	1.39	1.47	2	2
A	131	TYR	C-N	4.78	1.40	1.33	5	1
A	16	HIS	CG-ND1	4.78	1.43	1.38	1	2
A	112	HIS	CD2-NE2	4.77	1.43	1.37	1	1
A	126	GLN	C-N	4.76	1.40	1.33	4	1
A	157	ARG	CZ-NH1	4.76	1.26	1.32	9	1
A	163	GLU	N-CA	4.75	1.55	1.46	6	2
A	79	PRO	N-CD	4.75	1.41	1.47	6	1
A	111	LEU	C-N	4.75	1.40	1.33	5	1
A	183	ARG	CA-CB	4.75	1.62	1.53	6	2
A	105	PRO	C-N	4.73	1.26	1.33	2	1
A	152	PRO	C-N	4.72	1.40	1.33	10	1
A	131	TYR	N-CA	4.71	1.55	1.46	5	1
A	80	TYR	CG-CD1	4.71	1.49	1.39	3	1
A	171	LYS	CA-CB	4.70	1.62	1.53	1	1
A	139	TYR	C-N	4.66	1.26	1.33	7	1
A	152	PRO	N-CD	4.66	1.41	1.47	6	1
A	22	GLU	CA-CB	4.64	1.62	1.53	1	1
A	86	ALA	C-N	4.64	1.39	1.33	2	1
A	113	PRO	N-CD	4.63	1.41	1.47	4	3
A	39	LEU	C-N	4.63	1.39	1.33	1	1
A	151	ARG	CZ-NH1	4.60	1.26	1.32	4	1
A	179	ALA	C-N	4.57	1.39	1.33	7	1
A	163	GLU	CA-C	4.57	1.62	1.52	5	1
A	52	TYR	CB-CG	4.54	1.61	1.51	3	1

Standard geometry: angle outliers

There are 915 bond angle outliers in this entry (4.65% of 19680 assessed bonds). A summary is provided below. The output is limited to 100 rows.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	63	ASN	CA-CB-CG	10.20	122.80	112.60	3	3
A	124	HIS	CB-CG-CD2	9.93	118.29	131.20	4	6
A	159	GLN	OE1-CD-NE2	9.75	112.85	122.60	6	2
A	28	ARG	NE-CZ-NH2	9.51	127.75	119.20	1	3
A	142	ARG	NE-CZ-NH2	9.42	127.68	119.20	6	4
A	75	GLN	OE1-CD-NE2	9.24	113.36	122.60	8	6
A	119	PRO	N-CA-CB	9.12	113.03	103.00	1	6
A	6	HIS	CB-CG-CD2	9.06	119.43	131.20	2	5
A	166	ALA	C-CA-CB	9.01	124.01	110.50	5	1
A	37	ARG	NE-CZ-NH1	8.94	130.44	121.50	2	3
A	120	PHE	CA-CB-CG	8.79	122.59	113.80	7	4
A	92	ASN	CA-CB-CG	8.78	121.38	112.60	1	3
A	31	LEU	C-N-CA	8.69	137.35	121.70	3	3
A	146	PRO	N-CA-CB	8.67	112.54	103.00	6	2
A	140	THR	CA-CB-CG2	8.43	124.82	110.50	9	1
A	156	ASN	OD1-CG-ND2	8.36	114.24	122.60	2	4
A	49	PRO	N-CA-CB	8.36	112.19	103.00	4	3
A	38	PRO	N-CA-CB	8.33	112.16	103.00	7	6
A	101	ASN	OD1-CG-ND2	8.28	114.32	122.60	8	5
A	122	GLN	OE1-CD-NE2	8.23	114.37	122.60	9	6
A	17	GLN	OE1-CD-NE2	8.13	114.47	122.60	3	4
A	21	ASN	OD1-CG-ND2	8.11	114.49	122.60	3	2
A	99	PRO	N-CA-CB	7.97	111.77	103.00	7	3
A	30	GLN	OE1-CD-NE2	7.93	114.67	122.60	10	6
A	25	PRO	N-CA-CB	7.93	111.72	103.00	7	4
A	27	ASN	OD1-CG-ND2	7.88	114.72	122.60	2	4
A	16	HIS	O-C-N	7.78	110.55	123.00	2	1
A	55	PRO	N-CA-CB	7.77	111.55	103.00	7	4
A	112	HIS	CB-CG-CD2	7.74	121.13	131.20	2	4
A	175	ALA	N-CA-CB	7.73	98.80	110.40	4	3
A	79	PRO	N-CA-CB	7.72	111.49	103.00	7	5
A	126	GLN	CB-CG-CD	7.66	125.63	112.60	9	1
A	16	HIS	ND1-CE1-NE2	7.59	115.99	108.40	5	3
A	107	PRO	N-CA-CB	7.59	111.35	103.00	10	5
A	152	PRO	N-CA-CB	7.57	111.32	103.00	2	3

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	142	ARG	NH1-CZ-NH2	7.54	109.49	119.30	6	3
A	156	ASN	CA-CB-CG	7.46	120.06	112.60	2	2
A	71	GLN	OE1-CD-NE2	7.45	115.15	122.60	6	4
A	140	THR	OG1-CB-CG2	7.45	94.41	109.30	9	1
A	166	ALA	N-CA-CB	7.43	99.26	110.40	5	2
A	37	ARG	NE-CZ-NH2	7.41	112.53	119.20	2	2
A	62	PHE	CA-CB-CG	7.36	121.16	113.80	9	4
A	126	GLN	OE1-CD-NE2	7.25	115.35	122.60	6	5
A	56	GLU	C-N-CA	7.18	134.63	121.70	8	2
A	82	PRO	N-CA-CB	7.18	110.90	103.00	9	2
A	72	VAL	CA-CB-CG2	7.15	122.56	110.40	8	2
A	21	ASN	CA-CB-CG	7.11	105.49	112.60	7	2
A	158	ARG	NE-CZ-NH1	7.07	128.57	121.50	7	5
A	16	HIS	CB-CG-CD2	7.06	122.03	131.20	7	8
A	109	MET	C-N-CA	7.04	134.38	121.70	6	4
A	76	THR	CA-CB-CG2	7.04	122.46	110.50	3	2
A	153	ASN	OD1-CG-ND2	6.96	115.64	122.60	6	2
A	158	ARG	NH1-CZ-NH2	6.95	110.27	119.30	5	4
A	43	TYR	C-N-CA	6.92	134.15	121.70	6	1
A	34	PRO	N-CA-CB	6.88	110.56	103.00	5	4
A	97	PHE	CA-C-N	6.87	127.20	116.90	6	2
A	124	HIS	C-N-CA	6.86	134.05	121.70	4	1
A	45	ASP	CA-CB-CG	6.76	119.36	112.60	7	3
A	124	HIS	ND1-CG-CD2	6.74	112.84	106.10	7	2
A	34	PRO	C-N-CA	6.69	133.75	121.70	9	2
A	6	HIS	CG-CD2-NE2	6.68	100.52	107.20	4	2
A	47	SER	C-N-CA	6.66	133.69	121.70	6	1
A	6	HIS	O-C-N	6.65	112.36	123.00	1	1
A	47	SER	CA-CB-OG	6.65	124.40	111.10	2	1
A	29	PRO	N-CA-CB	6.64	110.31	103.00	4	5
A	112	HIS	CA-CB-CG	6.61	120.41	113.80	10	2
A	26	LEU	N-CA-CB	6.60	99.27	110.50	6	1
A	168	THR	C-N-CA	6.60	133.58	121.70	8	1
A	134	ASN	OD1-CG-ND2	6.60	116.00	122.60	10	3
A	53	ASN	OD1-CG-ND2	6.59	116.01	122.60	4	5

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	30	GLN	CG-CD-NE2	6.58	126.28	116.40	10	1
A	151	ARG	NE-CZ-NH2	6.58	113.28	119.20	3	3
A	86	ALA	N-CA-CB	6.52	100.62	110.40	2	1
A	124	HIS	ND1-CE1-NE2	6.50	114.90	108.40	10	3
A	164	ARG	NE-CZ-NH2	6.50	125.05	119.20	1	6
A	87	ALA	N-CA-CB	6.46	100.71	110.40	8	1
A	183	ARG	NE-CZ-NH1	6.45	115.05	121.50	1	3
A	112	HIS	CB-CG-ND1	6.44	132.36	122.70	2	1
A	63	ASN	OD1-CG-ND2	6.44	116.16	122.60	2	5
A	125	GLY	C-N-CA	6.42	133.26	121.70	5	2
A	171	LYS	C-N-CA	6.41	133.24	121.70	6	2
A	67	ALA	C-CA-CB	6.38	120.07	110.50	1	1
A	112	HIS	ND1-CE1-NE2	6.34	114.74	108.40	1	2
A	164	ARG	O-C-N	6.34	112.86	123.00	1	2
A	69	ASN	OD1-CG-ND2	6.32	116.28	122.60	6	2
A	6	HIS	CA-CB-CG	6.31	120.11	113.80	10	3
A	147	PRO	N-CA-CB	6.30	109.93	103.00	7	3
A	98	PRO	CA-C-N	6.27	126.31	116.90	1	2
A	177	GLU	CA-CB-CG	6.27	126.64	114.10	1	1
A	107	PRO	CA-N-CD	6.25	103.24	112.00	6	2
A	67	ALA	CA-C-O	6.22	131.38	120.80	8	1
A	24	GLU	CB-CG-CD	6.22	102.03	112.60	3	2
A	37	ARG	NH1-CZ-NH2	6.22	111.22	119.30	6	2
A	28	ARG	NE-CZ-NH1	6.19	127.69	121.50	9	3
A	162	ARG	C-N-CA	6.11	132.70	121.70	6	2
A	35	LEU	N-CA-CB	6.09	120.85	110.50	7	2
A	99	PRO	N-CD-CG	6.09	112.33	103.20	5	4
A	12	MET	C-CA-CB	6.07	121.64	110.10	6	2
A	76	THR	O-C-N	6.07	113.28	123.00	5	1
A	143	GLU	CB-CG-CD	6.05	102.31	112.60	9	1

Too-close contacts ?

The following all-atom clashscore is based on a MolProbity analysis. All-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The table below contains clashscores for all atomic models in this entry.

Model ID	Clash score	Number of clashes
1	0.00	0

Model ID	Clash score	Number of clashes
2	0.00	0
3	0.00	0
4	0.00	0
5	0.00	0
6	0.00	0
7	0.00	0
8	0.00	0
9	0.00	0
10	0.00	0

There are no too-close contacts.

Torsion angles: Protein backbone

In the following table, Ramachandran outliers are listed. The Analysed column shows the number of residues for which the backbone conformation was analysed.

Model ID	Analysed	Favored	Allowed	Outliers
1	182	147	32	3
2	182	169	10	3
3	182	154	26	2
4	182	161	14	7
5	182	153	23	6
6	182	159	19	4
7	182	153	23	6
8	182	156	19	7
9	182	155	23	4
10	182	150	28	4

There are 33 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	55	PRO	10
A	57	GLY	2
A	70	ALA	2
A	126	GLN	2
A	144	ALA	2
A	10	SER	1
A	11	GLY	1
A	14	LEU	1

Chain	Res	Type	Models (Total)
A	21	ASN	1
A	32	LYS	1
A	35	LEU	1
A	42	VAL	1
A	44	LEU	1
A	47	SER	1
A	69	ASN	1
A	72	VAL	1
A	74	GLY	1
A	76	THR	1
A	79	PRO	1
A	86	ALA	1
A	88	ALA	1
A	90	GLY	1
A	96	GLY	1
A	99	PRO	1
A	103	VAL	1
A	110	LEU	1
A	117	LEU	1
A	124	HIS	1
A	129	PRO	1
A	136	PRO	1
A	141	VAL	1
A	146	PRO	1
A	163	GLU	1

Torsion angles : Protein sidechains ?

In the following table, sidechain rotameric outliers are listed. The Analysed column shows the number of residues for which the sidechain conformation was analysed.

Model ID	Analysed	Favored	Allowed	Outliers
1	147	123	14	10
2	147	133	12	2
3	147	132	9	6
4	147	126	15	6
5	147	109	24	14

Model ID	Analysed	Favored	Allowed	Outliers
6	147	126	13	8
7	147	117	18	12
8	147	120	20	7
9	147	125	14	8
10	147	120	20	7

There are 63 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	140	THR	4
A	165	LEU	3
A	168	THR	3
A	4	THR	2
A	26	LEU	2
A	35	LEU	2
A	39	LEU	2
A	42	VAL	2
A	72	VAL	2
A	91	SER	2
A	97	PHE	2
A	126	GLN	2
A	177	GLU	2
A	1	MET	1
A	2	THR	1
A	6	HIS	1
A	12	MET	1
A	14	LEU	1
A	15	LEU	1
A	17	GLN	1
A	18	ILE	1
A	19	GLN	1
A	31	LEU	1
A	33	ILE	1
A	43	TYR	1
A	44	LEU	1
A	45	ASP	1

Chain	Res	Type	Models (Total)
A	51	VAL	1
A	52	TYR	1
A	60	TYR	1
A	78	LEU	1
A	85	GLU	1
A	99	PRO	1
A	100	LEU	1
A	102	SER	1
A	103	VAL	1
A	104	SER	1
A	106	SER	1
A	108	LEU	1
A	109	MET	1
A	110	LEU	1
A	112	HIS	1
A	116	GLN	1
A	117	LEU	1
A	120	PHE	1
A	127	GLN	1
A	128	VAL	1
A	129	PRO	1
A	132	LEU	1
A	134	ASN	1
A	143	GLU	1
A	151	ARG	1
A	154	SER	1
A	159	GLN	1
A	164	ARG	1
A	169	ASN	1
A	174	MET	1
A	176	MET	1
A	178	SER	1
A	180	LYS	1
A	181	GLU	1
A	183	ARG	1

Chain	Res	Type	Models (Total)
A	184	TYR	1

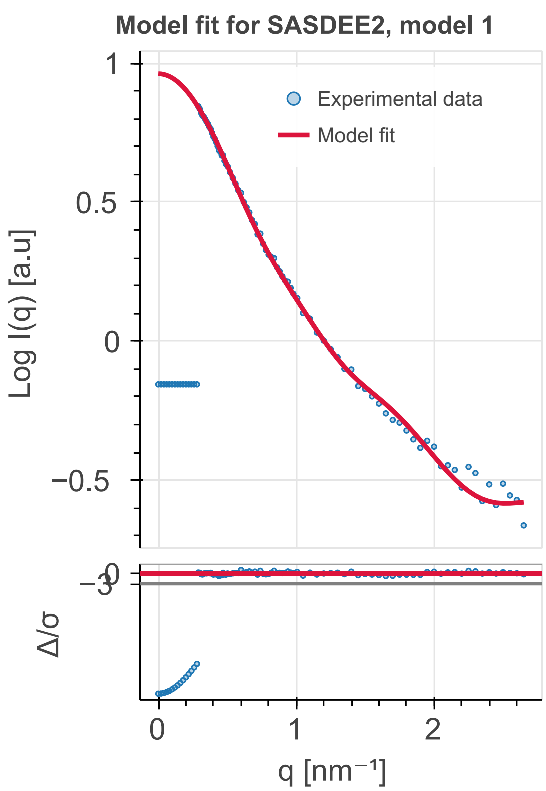
5. Fit to Data Used for Modeling Assessment

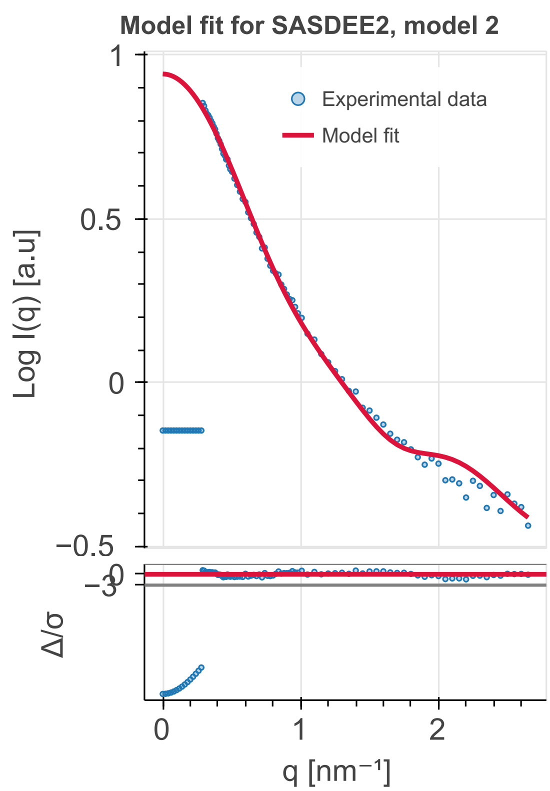
5.1. SAS

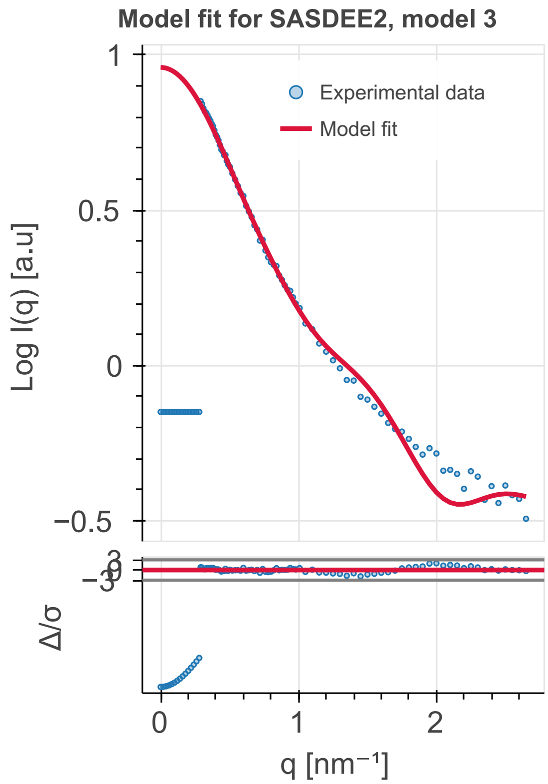
Model and fits displayed below were obtained from SASBDB.

5.1.1 Model versus Experimental Scattering Profiles

Experimental (blue) and model (red) scattering profiles are presented as log *I(q) vs. q* together with error weighted residual difference plot between the experimental and model *I(q) vs q*. The *I(q)* is the intensity (preferably in absolute units of cm-1 or arbitrary units) and *q* is the modulus of the scattering vector. For a good fit, residual values should be equally and randomly spaced around the horizontal axis largely within +/- three standard deviations.







5.1.2 χ^2 goodness of fit and 5.1.3 cormap analysis

χ^2 values are a measure of the overall fit of the model to the 1D scattering profile. A model that fits the data within its error estimates will have a χ^2 value close to one, provided that the dominant errors are the random statistical errors (i.e. no systematic errors) from the SAS measurement that are correctly propagated. Correlation Map (CorMap) test is a variance-covariance analysis on the scattering intensities comparing two (or more) scattering profiles (e.g. model versus experiment or multiple measures from the same sample). The CorMap test complements χ^2 and importantly is independent of the reported errors. The method assigns a probability (P-value based on a 1-tailed Schilling test) for finding the longest string of experimental data points that lie systematically above (+1) or below (-1) the model profile. The P-value lies between 0 – 1 and a significance threshold is chosen below which the model fit is judged to show systematic deviation from experiment. A typical range statisticians use to indicate significant deviation is 0.01 - 0.05. As implemented in the ATSAS suite, the reported CorMap P-value is green (model fit is good) for $P > 0.05$, yellow for $0.01 < P < 0.05$, and red (model deviates significantly) for $P < 0.01$.

SASDB ID	Model	χ^2	P-value
SASDEE2	1	0.14	0.00
SASDEE2	2	0.34	0.00
SASDEE2	3	0.51	0.00

5.4. Hydroxyl radical footprinting

Validation for this section is under development.

6. Fit to Data Used for Validation Assessment

Validation for this section is under development.

Acknowledgments

The development of integrative model validation metrics, implementation of a model validation pipeline, and creation of a validation report for integrative structures are funded by NSF awards to the PDB-IHM team (DBI-1756248, DBI-2112966, DBI-2112967, DBI-2112968, and DBI-1756250) and awards from NSF, NIH, and DOE to the RCSB PDB (DBI-2321666, R01GM157729, and DE-SC0019749). The PDB-IHM team and members of the Sali lab contributed model validation metrics and software packages.

Dr. Jill Trewhella, Dr. Dina Schneidman, and members of the SASBDB repository are acknowledged for their advice and support in implementing SAS validation methods. Team members from the labs of Dr. Juri Rappsilber, Dr. Alexander Leitner, Dr. Andrea Graziadei, and members of PRIDE database are acknowledged for their advice and support in implementing crosslinking-MS validation methods. We are grateful to Dr. Shruthi Viswanath for discussions about uncertainty assessment of integrative structural models.

Members of the wwPDB Integrative/Hybrid Methods Task Force provided recommendations and community support for the project.