

**Supplementary Material for:**

**A fundamental catalytic difference between Zinc and Manganese dependent enzymes revealed in a bacterial Isatin hydrolase**

**Theis Sommer<sup>1,\*</sup>, Kaare Bjerregaard-Andersen<sup>2,\*</sup>, Lalita Uribe<sup>5,6\*</sup>, Michael Etzerodt<sup>4</sup>, Gregor Diezemann<sup>5</sup>, Jürgen Gauss<sup>5</sup>, Michele Cascella<sup>2,7</sup>, J. Preben Morth<sup>1,3</sup>**

<sup>1</sup> Norwegian Center for Molecular Medicine, Nordic EMBL Partnership University of Oslo, Gaustadalléen 21, 0349 Oslo, Norway

<sup>2</sup> Department of Chemistry, University of Oslo, Sem Sælands vei 26, 0371, Oslo, Norway

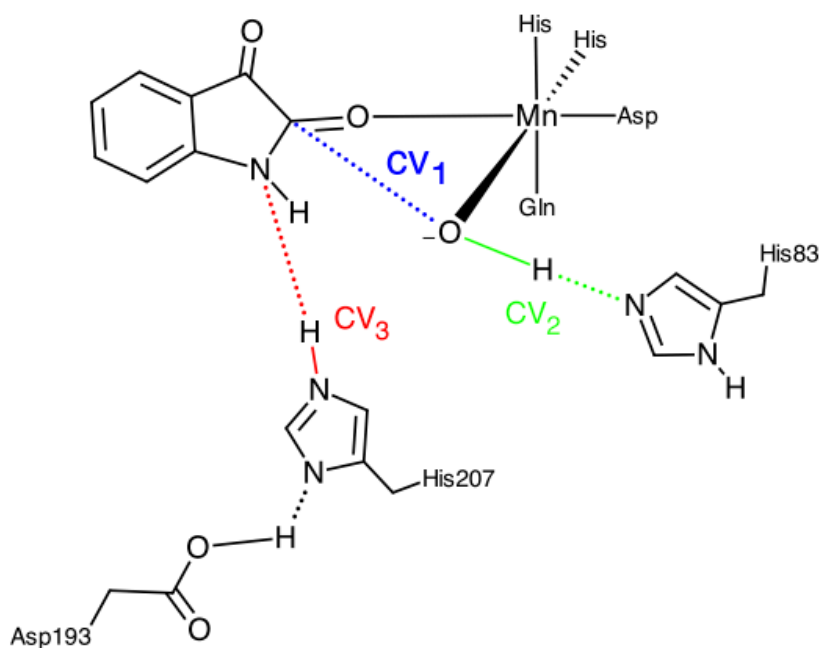
<sup>3</sup> Institute for Experimental Medical Research, Oslo University Hospital, N-0424 Oslo, Norway

<sup>4</sup> Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds vej 10C, DK-8000 Aarhus, Denmark

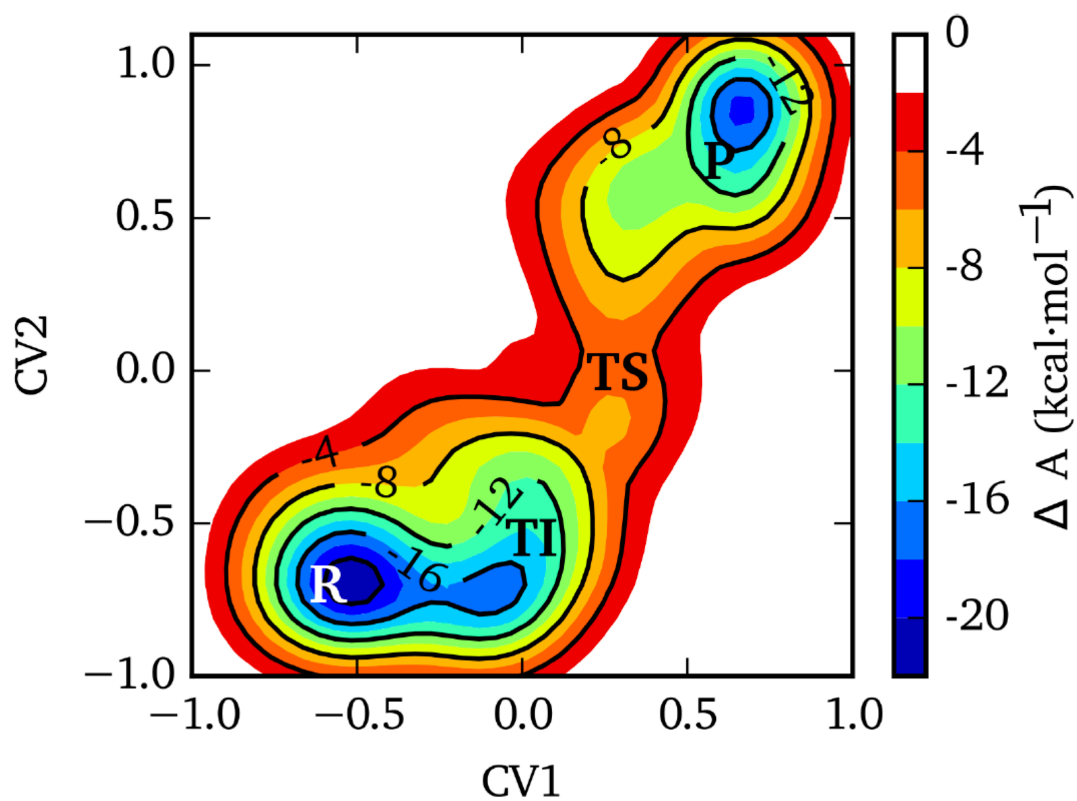
<sup>5</sup> Institut für Physikalische Chemie, Johannes-Gutenberg Universität Mainz, Duesbergweg 10-14, 55128 Mainz, Germany

<sup>6</sup> Graduate School Materials Science in Mainz, Staudinger Weg 9, 55128, Germany

<sup>7</sup> Hylleraas Centre for Quantum Molecular Sciences, University of Oslo, Sem Sælands vei 26, 0371, Oslo, Norway



Supplementary Figure S1, Graphical view of the collective variables (CV<sub>1</sub>-CV<sub>3</sub>) defined in the metal binding pocket of *laIHA*, for the QM/MM - metadynamics simulations presented in the main text.



Supplementary Figure S2, Projection of the free energy surface over the CV1-CV2 coordinates. The letters in the plot indicate: (R) the reactant state, (TI) the tetrahedral intermediate, (TS) the transition state, and (P) the isatinate product.

a

Species	Accession	Length	Sequence	Position
<i>LaIHA</i>	(5nna)	1	-----MSS-----LNQLVSLGASGAVRVDLTHLTDLPDPFVIVLP-----PEFGQCAR-FRMEETISAYDHRGPAWKWHNIS	65
<i>LaIHB</i>	(4j0n)	1	-----MSAQS-----ALSLGAKLSEGEVVDVDTQVGLGNPTPIQLP-----PDFAKNTPKVEIHKISEYSDSGPFFAWNMV	69
<i>RsIHA</i>	(5nmp)	1	-----MTSTTPTILPALAAGLARGNIRVDLTQTLSPSFPTLQLP-----SQFGQVQP-FKIERISHYDASGPAWVWNFS	70
<i>BcKynB</i>	(4coc)	1	-----MDTL-----WDISPPVSPATPVWPG-----DTPVAVERVWRMEAGSPVNVVA-RLT	44
<i>PaKynB</i>	(4coc)	1	-----MTSLRYWDISPALDPNTPTWPG-----DTPFQEQWAARLDEQCPVNVG-RIT	46
<i>BaKynB</i>	(4coc)	1	-----MKTSLWIDISQPLNDIATWPG-----DTPFSEYVLSKSESGSNVVG-KLT	46
<i>BsAHD</i>	(1r61)	1	-----AAMKVYDVTAPIYEGMPVYKN-----KPEKQPKRTTITNGVYTESRID	43
<i>MjAHD</i>	(2b0A)	1	-----MTLLDLDLQTLIN-FPYPGDF-----ELRILEKKIDGFIIVSEII	37
<i>RzHpoH</i>		1	MTDLTISAAGLRADIAVATSPYGPEDERGLNKKITPESRAALSRVDSGRVYDLSIDYFVGMPSFQAAGDPGYQIFMSHTPAGTAVDNLNGVGEAVNRHVCYSGDVVF	109
<i>LaIHA</i>	(5nna)	66	MSEHTCTHFDAPSHWISGKDVNGSVDIEPAEA-----FVGPVVVDCSKGA-----AENDDFELTPEIIAGWESEHGRIPEDAWVLMRTDWSKRRG--ADY	155
<i>LaIHB</i>	(4j0n)	70	LGEHSTHFDAPHHWITGKDYSDGFTDLDVQR-----LIAPVNVDCSKES-----AADPDFLLTADLIKAWAEHGEIGAGEVVMRTDWDKRGDEAAF	161
<i>RsIHA</i>	(5nmp)	71	CGEHTCTHFDAPAHWITGRDYPGNSVDTIAPEN-----FVAPAVVIDASAQV-----RENEDWLLTVDFLQAEQRHGRIPAGAVLFRDWSLRVGDAAAF	162
<i>BcKynB</i>	(4coc)	45	LSPHGTGACDAPLHYRAD---GAPIGAVPLDT-----YLGPCRVIHICIG-----AAPVVRPADVEAALDGVPPRVLLTYARAARVQWDSN	122
<i>PaKynB</i>	(4coc)	47	LSPHGTGAVDGPLHYRAD---GLPIGQVPLDI-----YMGPCRVIHICIG-----ANPLVTPHELAGQLDLDLPSRVLLRTERFVNPAN-WPEG	123
<i>BaKynB</i>	(4coc)	47	MSIHTCTHIDAPLPHDND---GKKVLDLDLQIV-----YVGPTRIIDVSN-----LESIGKKELEKPHLEGV-ERLLLTSSHGKANEPDI	123
<i>BsAHD</i>	(1r61)	44	MVHTCTHIDAPLHVEG---GATFETIPLND-----LVGPKLFDLTH-----VNDIRTKDIIAHLDIQEGDFVLPYKTKNSFEDAFHFF	121
<i>MjAHD</i>	(2b0A)	38	MSEHSLCTHIDYKIVGLEN---RIPFKDG-----IIGKGG-----YCISLDDFERNKLPACDILLITGFSKYWGRDEYF	104
<i>RzHpoH</i>		110	MHTHTCTHIDALNIFGVDGKTIYNNFTVEENLGRHWKGGAEITPIVARGVLLDIALKGVLECLPPSYAITVEDCQAALKEAGTELRDGDVALITRGMRYVDPDGSKV	218
<i>LaIHA</i>	(5nna)	156	LNMADGPHSPGTPPEAIRFLIEERNIRGFTTETVGT*AGQGAHYVPPYPAHYLLHGAGKYGLQC LANLDQLP---ATGAVLIAAALPKIKNGTGSLEVLAMVTE--- 257	
<i>LaIHB</i>	(4j0n)	162	LNADETPHSPGTPDAEYLLS-KKIVGWSSQICIGT*AGQAGGMEPPFPAHNLHRDNCFLASLANLDKLP---AKGAILIAAALPKIKNGTGSLEVLAMVTE--- 263	
<i>RsIHA</i>	(5nmp)	163	LNIREDGHTPGTQEAWEVLIGERNVHGFSVETINT*AGQSYAWPLPCHTMLHGANNRYGLQCLKNLDQLP---PRGAFILAAALPKIEGGSSGLVNLALVE--- 263	
<i>BcKynB</i>	(4coc)	123	FCVAVP-----DVTDLAAGVGLIGIDTPLDLPQES---KTMADARRVRAHRMALLEGIV-LDDVP---PGDYELIALPKLFTKFDASVRAVRLALPAQAS 213	
<i>PaKynB</i>	(4coc)	124	FCATAP-----ATIECLAERGVRLVIGIDTPLDLPQES---KTLDAHRAVGRHGMALLEGIV-LDDVP---AGDYELIALPKLFTKFDASVRAVRLALPTAE 213	
<i>BaKynB</i>	(4coc)	124	IPHLRA-----DIAPFLSEKGIKLVGDDVPSVPLD---KELAAHQKLFKHSIHLENNV-LDHVA---DGDYELIALPKLFTKFDASVRAVRLALPTAE 209	
<i>BsAHD</i>	(1r61)	122	IFVAED-----AARYLADKQIRGVGIDALGTERAQEG---HPTKTLIFSAGVILLEGRL-LKDVV---EGRYFMVAALPKLVGTDAAPARVLLDFREP--- 207	
<i>MjAHD</i>	(2b0A)	105	EKIPETIP-----FLDDIKISNICKVGDICTIG---GFPEEKRLLSNNILITLNLN---ENLKNLVGKSFYFLGLPKLFDIDASPTKCAIIL--- 186	
<i>RzHpoH</i>		219	LGNPPPLG-----LDAARWITAQGVAVVADQCEV*VGPSEHDDNWLPGHCFELAEAGVPMIELVNLLEELARDEVHEFCLMAA*IKLRGASGATLEPLAMPRLTGDR 320	

b

Species	Accession	Length	Sequence	Position
<i>EcAAH</i>	(4pxd)	1	-----MITHFRQAIETLPWL-SSFGADPAGG-----MTHLLYSP	34
<i>AtAAH</i>	(4pxb)	1	MESLKRFLCSIALLLISLPPSSLAQQQHSIRTMEDFSGYPIHEFGQFPGSINLASSLSDVAPGLQNIQIDEL-SSFSADSPSS-----VTVLTYTD	91
<i>BvHYD</i>	(5i4m)	1	-----MDAVSETAKRAALDTSIKVDGRRLWDSLMEV-AKIGATPKGG-----VCRALALTD	49
<i>BmHYD</i>	(5thw)	1	-----MNAVSEALKRATIDPSIKVDGKRLWLSMEM-AKIGATPKGG-----VCRALALTD	49
<i>BcHYD</i>	(4wjB)	1	-----MRDAPASMPRVDDGRDLWASLERM-AQIGATPKGG-----VCRALALTD	41
<i>GeRAC</i>	(3n5f)	1	-----MIQGERLWQRLMEL-GEVGKQPSGG-----VTRLSFTA	32
<i>SkSYN</i>	(2v11)	1	-----MSKDVSSITTTVSASDPGTLNLPAAAPLSIASGRINQTILETGSQLGGVARWQESHEFGMRRLAGTA	68
<i>EcAAH</i>	(4pxd)	35	EWLETQQQFKRMAASLETRFDEVGNLYGRNLCTEYQPEVLLSGSHIDTVVNGNLDDQFQALAAWLAIDWLKTQYGAPLRTEVVMAMAEESGSRTPY	133
<i>AtAAH</i>	(4pxb)	92	KDVSARRVYVKNMLMALAGLTVREDAVGNIFGKWDGLEPNLPAVATGSHIDTVPYSGKYDGVVGVIGAIENLVNKRSGFKPKRSLEIILFTSEPTFRGI	190
<i>BvHYD</i>	(5i4m)	50	LDKAARDLIVGWAKAAACTVTVTDMGNVFMRRAGRVADAAPVVTGSHADSOPTGRFDGIYGVILGGLVIRSLNDHGIETEPHEVVIWVWNEGSSRAP	148
<i>BmHYD</i>	(5thw)	50	LDKAARDLIVGWAKAAACTVTVTDMGNVFMRRAGRVADAAPVVTGSHADSOPTGRFDGIYGVILGGLVIRSLNDHGIETEPHEVVIWVWNEGSSRAP	148
<i>BcHYD</i>	(4wjB)	42	LDRESRDLFVQWAREAACTVTVTDMGNVFMRRAGRVADAAPVVTGSHADSOPTGRFDGIYGVILGGLVIRSLNDHGIETEPHEVVIWVWNEGSSRAP	140
<i>GeRAC</i>	(3n5f)	33	EERRAKDLVASYMREALFVYEDAAENLIGRKEETNPDATVVLVSHLSDVYNGCFDGLGLVAGVEVVQTMNEHGVVTHHPTEVVAFTDEEGARFP	131
<i>SkSYN</i>	(2v11)	69	LDGAMRDFWTNECESLCKRVKIKIENMFAVYFK-NGGKPTATGSHLDTPQEAQYDGLVILGVILAGLEVLRTEKDNVYVPIVQVWVWNEGSSRAP	166
<i>EcAAH</i>	(4pxd)	134	VFWGSKNIFGLANPDVVRN-ICDAKG--NSFVDMACKCFLLPN----APLTPRQDIKAFVLEHIEQCVLESNGQSIQVNVAINVQRRYTVTLNDES	224
<i>AtAAH</i>	(4pxb)	191	SCLSGRLLAGSKELAEALK-TTVVDGQNVSFIEAARSAEYEDKDDLSVFLKKGSYFAPLELHIEQCPILDEDEGLDGVVTAIAAPASLKVPEFENG	288
<i>BvHYD</i>	(5i4m)	149	AMVASGVFAGVFPLEYG--LSRKVDVGKTIIEGELARIYAGDAP-----CGGRK--LHAFAELHIEQCPILAEAEKTIQVVTAGQQRWYELTPTQDE	237
<i>BmHYD</i>	(5thw)	149	AMVASGVFAGVFPLEYG--LSRKVDVGKTIIEGELARIYAGDAP-----CGGRK--LHAFAELHIEQCPILAEAEKTIQVVTAGQQRWYELTPTQDE	237
<i>BcHYD</i>	(4wjB)	141	AMVASGVFAGVFPLEYG--LSRVDGAGRTIEGELERISYAGAEV-----VGGYP--VHAAYELHIEQCPILAEAEKTIQVVTAGQQRWYELTPTQDE	229
<i>GeRAC</i>	(3n5f)	132	MGISGRAMAG-TLPPALE-CDRAEG--ISLAEMAKQLDPRD---LPQAARKPGTVKAVVELHIEQCVLEETGLPVGIIVTGTAGLIVKVFTEEKA	223
<i>SkSYN</i>	(2v11)	167	SGTSSVSWSHDLSLEAYGLMSVGEOKPEVSYDLSKNIYIGIDTP-----ASYKENEIDAFHLEHIEQCPILAEENKATIGITVGVQAYNQKRVTVHVG	260
<i>EcAAH</i>	(4pxd)	225	NHAGTTPMGRYRITVYAFSRICHSVEKAKRMG-DPLVLTDFQKVEPRNTVNVV*GKTFITIDCRHTDAAVLRDFT---QOLENDMRAICDEMIGIDI	319
<i>AtAAH</i>	(4pxb)	289	GHAGAVLMPYRNDAGLAAELALAVEKHVLESSEIDTVGVIGILELHGAINSIKSKHLEIDTRIDEARRNTVI---KKIQESANTIAKKRKKVLS	384
<i>BvHYD</i>	(5i4m)	238	AHAGTTPMRRRALLGASRVVDLVNRIGLDHAP-YGCAIVGMQVHNSRNVIGRVFETVDFRHPDDAVLAKMD---AALRDGVARIADIDGLDAL	332
<i>BmHYD</i>	(5thw)	238	AHAGTTPMRRRALLGASRVVDLVNRIGLDHAP-YGCAIVGMQVHNSRNVIGRVFETVDFRHPDDAVLAKMD---AALRDGVARIADIDGLDAL	332
<i>BcHYD</i>	(4wjB)	230	AHAGTTPMARRRDLVGAARMISFVEVLGRRYAP-DARATVGMIEARNSRNVIGCGCFETVEFRHPDDAVLDELDD---AALRAELARVADETLGQAQI	324
<i>GeRAC</i>	(3n5f)	224	EHAGTTPMSLRDPMMAAAQIIIVIEEAEARTG--TVGVGLQHVYGGIIVIERVEFVLDLRLKAEVRDQVW---KAIAVRAETIAKERNVYTT	317
<i>SkSYN</i>	(2v11)	216	AHAGTTPMRLKALLMSSKMIVAASEIAQRHN---GLFPCITIDAKYYSVNIIGEVSF*LDLFRHPSDDVLAATMLKEAAEFDRILIKINDGALSYS	356
<i>EcAAH</i>	(4pxd)	320	DLWMEDEEVPMM-KELVATLTLCERK-LNYRVMHSGHDAQIFRPRVTCMI*IPISINQISHNPA*RTNITDLAEIVKTI*ALMLYQLAWQK----	411
<i>AtAAH</i>	(4pxb)	385	FKIVNQDFPALSCLKVKKMAEAATELN-LSHKMMIERAYHDSLFMRIRSMGMIIFPCYKGYSHKPEYSSPEDMANGVKVSILTAKLSLD-----	476
<i>BvHYD</i>	(5i4m)	333	EQIFYYAPAFDFAACVAV-RAAADRFY-YSHRDIVSGAGHDACYLAQVAPTSVVFVPCDISHNEI*DATPAWIEAGANVLLHAML*SRACEPVS---	426
<i>BmHYD</i>	(5thw)	333	EQIFYYAPAFDFAACVAV-REAAERFG-YPHRDIVSGAGHDACYLAQVAPTSVVFVPCDISHNEI*DATPAWIEAGANVLLHAML*SRACEPVS---	426
<i>BcHYD</i>	(4wjB)	325	EQIFYYAPVFPFRCIDTV-RDAARSLG-LSHMDIVSGAGHDACYVAVAPTMGIFVPCDGLSHNEA*ATPEWATAGADVLLRAVL*SAQEA-----	416
<i>GeRAC</i>	(3n5f)	318	ERLQ-EMPPVCSDEVKRAEAACQKLG-YPSFWLPSGAHDSVQLAPIC*IGMIFVRSQDGVSHSPA*WSTKEDCAA*AEVYHT*VWQLAQE-----	409
<i>SkSYN</i>	(2v11)	357	ETLQVSPAVNFHEVCEICVSRSAFAQFKKQDQVRQI*WAGHDSQCT*PHVTSMT*ITP*SKDGLSHNY*YSSPEEIE*NFVLLQAI*INDYNRVIRGH	455

Supplementary Figure S3. Complete sequence alignments of partial alignment shown in Figure 2a. a) Isatin hydrolase homologues, b) Allantoate amidohydrolase homologues. Conserved residues are highlighted in blue. Residues involved in metal coordination are indicated by a dot, residues presumed to be involved in proton transfer are indicated with an asterisk. The key GLQC/GLAS motif of IHA and IHB, indicating isatin hydrolase activity, is highlighted in green, with a similarly conserved motif in KynB homologues highlighted in red.

**Supplementary Table 1. X-ray crystallographic data collection and refinement statistics for IHA**

	<i>LaIHA:Benzylozoate</i> <sup>1</sup>	<i>RsiHA</i> <sup>1</sup>	<i>LaIHA:isatin</i> <sup>1</sup>
PDB ID	5NNA	5NMP	5NNB
<b>Data collection</b>	Diamond, I02-P1	ESRF, ID30B	PETRA III
Space group	P 1	P 32 2 1	P 1
Wavelength (Å)	1.0004	0.9762	1.033
Cell dimensions <i>a, b, c</i> (Å)	37.7, 70.7, 92.9	164.8, 164.8, 275.9	38.4, 69.9, 93.5
$\alpha, \beta, \gamma$ (°)	108.2, 95.9, 101.5	90.0, 90.0, 120.0	108.2, 95.3, 103.8
Resolution (Å) <sup>2</sup>	86.86 – 1.50 (1.54 – 1.50)	99.2 - 2.65 (2.65-2.72)	19.5 - 1.79 (1.9 – 1.79)
$R_{\text{merge}}^2$	0.073 (0.421)	0.093 (1.093)	0.052 (0.496)
$I / \sigma I^2$	10.1 (2.0)	15.6 (1.5)	7.99 (1.0)
Completeness (%) <sup>2</sup>	96.1 (93.7)	99.8 (99.4)	94.05 (94.6)
Redundancy	3.4	6.9	1.8
CC <sub>1/2</sub>	0.995(0.937)	0.997 (0.658)	0.996 (0.816)
<b>Model refinement</b>			
No. observed reflections <sup>2</sup>	461622 (32325)	867299 (64074)	263115 (41566)
Unique <sup>2</sup>	135172(9712)	125897 (9226)	141099 (22382)
$R_{\text{work}}$	0.149	0.152	0.1826
$R_{\text{free}}$	0.180	0.196	0.2332
No. atoms	9195	20605	8307
Protein	8065	20300	784
Ligand/ion	106	10	48
Water	1024	295	419
$B$ -factors (Å <sup>2</sup> )			
Protein	12.2	68.8	40.5
Water	26.9	57.0	54.0
Ligand/ion	21.1	58.6	41.5
Monomers (asymmetric unit)	4	10	4
R.m.s. deviations			
Bond lengths (Å)	0.006	0.008	0.007
Bond angles (°)	1.06	1.04	1.01
Ramachandran plot, residues in:			
Favored region (%)	98	96	98
Allowed regions (%)	1.6	3.5	1.6
Outlier region (%)	0.4	0.5	0.4

<sup>1</sup> Complete diffraction data was obtained from one crystal

<sup>2</sup> Values in parentheses are for highest-resolution shell.