BIOLOGICAL
CRYSTALLOGRAPHY

Volume 70 (2014)

Supporting information for article:

Elucidation of the bicarbonate binding site and insights into the carboxylation mechanism of (*N*^{5})-carboxyaminoimidazole ribonucleotide synthase (PurK) of *Bacillus anthracis*

Micheal L. Tuntland, Bernard D. Santarsiero, Michael E. Johnson and Leslie W.-M. Fung

Table S1:

Table of ligand atom contacts and distances for ligands in 3QFF, 3R5H and 3V4S.

		3QFF (ADP)				3R5H (ADP)				3V4S (ADP/ATP)			
		ChA (Active Site 1)		**ChB (Active Site 2)**		**ChA (Active Site 3)**		**ChB (Active Site 4)**		**ChA (Active Site 5)**		**ChB (Active Site 6)**	
ADP/ATP	N1	V185(N)	2.7	V185(N)	2.8	V185(N)	2.8	V185(N)	2.8	V185(N)	2.7	V185(N)	2.9
	N3					H2O 639	3.0	H2O 403	3.0	H2O 390	3.0		
	N6					E182(OE2)	2.8	E182(OE2)	3.1	E182(OE2)	2.9	E182(OE2)	3.0
		K183(O)	2.8	K183(O)	3.0	K183(O)	2.8	K183(O)	3.1	K183(O)	3.0	K183(O)	3.0
	N7												
	N9												
	O2'			E190(OE1)	2.7	E190(OE1)	2.8	E190(OE1)	2.7	H2O 390	3.0	E190(OE1)	2.6
	O3'	E190(OE1)	3.1	E190(OE2)	2.8	E190(OE2)	3.0	E190(OE2)	2.7	E190(OE2)	2.8	E190(OE2)	2.6
						H2O 656	3.0	H2O 571	3.2				
	O4'			N216(ND2)	3.3							N216(ND2)	3.1
	O5'			Q158(OE1)	3.2							H2O 355	3.0
	O1A	H2O 509	2.7			Mg^{2+}	2.1	Mg^{2+}	2.4	Mg^{2+}	2.1	Mg^{2+}	2.1
				K183(O)	3.0	H2O 579	3.3	H2O 571	3.1	H2O 370	3.1	H2O 268	3.0
		N267(ND2)	2.9	N267(ND2)	3.0	N267(ND2)	3.2	N267(ND2)	3.1	N267(ND2)	3.2	N267(ND2)	3.2
						H2O 656	3.1						
	O2A					R107(NH2)	3.0			R107(NH2)	3.1		
		K147(NZ)	2.8	K147(NZ)	2.8	K147(NZ)	2.7	K147(NZ)	2.6	K147(NZ)	2.8	K147(NZ)	2.8
												Q158(NE2)	3.0
												E268(OE2)	2.9
	O3A			R107(NH2)	2.7	R107(NH1)	3.0						
						H2O 526	2.9						
	O1B					H2O 526	3.1	H2O 473	2.6			Y153(N)	3.3
				E255(OE1)	2.7					G155(N)	2.8	G155(N)	2.8
				E268(OE1)	2.9			E268(OE1)	2.9	N216(ND2)	3.1		
								E268(OE2)	2.9				
	O2B	R107(NH1)	2.7	E255(OE2)	3.0			H2O 473	3.1	R107(NH1)	2.5	R107(NH2)	2.7
		E268(OE1)	3.0	E268(OE2)	3.1	E268(OE1)	2.9			E268(OE2)	3.0	E268(OE1)	3.0
												H2O 101	2.2
												H2O 129	3.1
	O3B	H2O 509	2.6			Mg^{2+}	2.7	Mg^{2+}	3.2	Mg^{2+}	2.1	H2O 368	3.2
										H2O 410	3.0		
				H213(NE2)	2.8	N216(OD1)	3.0	R107(NH2)	3.0	E268(OE1)	3.3	H2O 101	3.4
						H2O 656	2.7						
	O1G											E255(OE2)	3.2
												Mg^{2+}	2.2
												H2O 368	3.1
	O2G											D154(N)	2.8
												H2O 368	2.9
												H2O 375	2.7
	O3G											H2O 101	2.0
												H2O 220	3.3
		3QFF (ADP)				**3R5H (ADP)**				**3V4S (ADP/ATP)**			
		ChA (Active Site 1)		**ChB (Active Site 2)**		**ChA (Active Site 3)**		**ChB (Active Site 4)**		**ChA (Active Site 5)**		**ChB (Active Site 6)**	
Mg^{2+}						H2O 656	2.7	H2O 473	3.1	H2O 370	2.2	H2O 268	2.0
						H2O 579	2.7	H2O 571	2.5	H2O 410	2.7		
						H2O 387	2.2	H2O 611	2.6				
						H2O 466	3.0						
								E255(OE1)	2.6			E255(OE1)	2.3
						E268(OE1)	2.2	E268(OE1)	2.2	E268(OE1)	2.0	E268(OE1)	2.2
						ADP (O1A)	2.1	ADP (O1A)	2.4	ADP (O1A)	2.1	ATP (O1A)	2.1
						ADP (O3B)	2.7	ADP (O3B)	3.2	ADP (O3B)	2.1	ATP (O1G)	2.2
Mg^{2+} at L269												H2O 183	
		3QFF (ADP)				**3R5H (ADP)**				**3V4S (ADP/ATP)**			
		ChA (Active Site 1)		**ChB (Active Site 2)**		**ChA (Active Site 3)**		**ChB (Active Site 4)**		**ChA (Active Site 5)**		**ChB (Active Site 6)**	
AIR	N3					Y153(OH)	3.0	H2O 576	2.7				
	N6					H2O 388	2.7	H2O 479	2.6				
						H2O 658	2.5	H2O 604	3.0				
	O2'					E76(OE1)	2.5	E76(OE1)	2.5				
						L19(N)	3.1						
	O3'					E76(OE2)	2.5	E76(OE2)	2.6				
						H2O 550	3.5	H2O 451	3.2				
								H2O 608	3.1				
	O4'					Q18(OE1)	3.4	H2O 394	2.9				
	O5'					K340(NZ)	3.3	H2O 394	3.1				
						H2O 395	2.9						
	O6					H2O 629	3.3	R347(NH2)	3.0				
								H2O 432	2.7				
								H2O 578	2.4				
	O7					R347(NH2)	3.1	K340(NZ)	2.9				
								R347(NE)	2.7				
						H2O 599	3.0	H2O 424	2.8				
	O8					R347(NE)	2.5						
						H2O 463	3.2						
		3QFF (ADP)				**3R5H (ADP)**				**3V4S (ADP/ATP)**			
		ChA (Active Site 1)		**ChB (Active Site 2)**		**ChA (Active Site 3)**		**ChB (Active Site 4)**		**ChA (Active Site 5)**		**ChB (Active Site 6)**	
HCO_3^-	O1									R272(NH2)	2.8	R272(NH2)	2.8
										H274(NE2)	3.3	H274(NE2)	3.4
												H2O 375	3.0
	O2									R272(NH1)	2.8	R272(NH1)	2.7
										H2O 12	2.8	H2O 41	3.1
	O3									K348(NZ)	3.4	K348(NZ)	3.4
										H274(NE2)	3.4	H274(NE2)	3.4

*Distances in Angstroms (Å)

*H2O in RED believed to be involved in transfer in our proposed mechanism

Figure S1:

Alignment of *ba*PurK amino acid sequence to that of 11 other PurKs for differing organisms; including from Archaea (*Sulfolobus solfataricus*), Bacteria (Gram-negative; *Escherichia coli*, *Vibrio cholerae*, *Aquifex aeolicus*; Gram-positive; *Bacillus subtilis*, *Staphylococcus aureus*; No Gram-stain; *Mycobacterium leprae*, *Mycobacterium tuberculosis*), Fungi (*Aspergillus clavatus*, *Saccharomyces cerevisiae*) and Plantae (*Vigna aconitifolia*). White letters in black boxes are strictly conserved residues throughout the 12 organisms. Loops are boxed and color-coded as in Fig 1A of text. A-Loop (green), B-Loop (blue), ABC-Loop (orange), DT-Loop (purple), J-Loop (cyan), P-Loop (red) and Ω-Loop (yellow).

```
B.anthracis       1 --------------MTRIILPGKTIGIIGGGQLGRMMALAAKEMGYKIAVLDPTKNSPCAQV-----ADIE 52
A.clavatus        1 ------------------MWNSRKVGVLGGGQLGRMLVESANRLNIQVNVLDAD-NSPAKQ---ISAHDGH 49
S.cerevisiae      1 -------------------MDSRTVGILGGGQLGRMIVEAANRLNIKTVILDAE-NSPAKQ---ISNSNDH 48
E.coli            1 ---------------------MKQVCVLGNGQLGRMLRQAGEPLGIAVWPVGLDAE-PAAVP-----FQQS 44
V.cholerae        1 ----------------------MRVLVLGAGQLARMMSLAGAPLNIETIAFDVGSE-NIVHP-----LTQT 43
A.aeolicus        1 ---------------------MLTVGILGGGQLGWMTILEGRKLGFKFHVLEDKENAPACRV-----ADRC 45
M.leprae          1 MMAVPSR---------CSLGVAPLVAMVGGGQLARMTHQAAIALGQTLRVLATAADEPAAQV-----TPDV 57
M.tuberculosis    1 MMAVASSR-----TPAVTSFIAPLVAMVGGGQLARMTHQAAIALGQNLRVLVTSADDPAAQV-----TPNV 61
B.subtilis        1 -------------MSKQIIYPGAVIGIIGGGQLGKMMAVSAKQMGYKVAVVDPVKDSPCGQV-----ADVE 53
S.aureus          1 -------------MNFNKLKFGATIGIIGGGQLGKMMAQSAQKMGYKVAVLDPAEDCPCRYV-----AHEF 53
S.solfataricus    1 --------------MFSVLDWKPKIGILGGGQLGWMIVLEGRKYPFTFYVLENDKNAPACRI-----ADRC 52
V.aconitifolia    1 ------------------GLYEVVVGVLGGGQLGRMMCQAASQMAIKVMVLDPQENCPASS-----LSYHH 48

B.anthracis      53 IVASYDDLKAIQHLAEISDVVTYEFENIDYRCLQWLEKH-A--YLPQGSQLLSKTQNRFTEKNAIEKAGLP 120
A.clavatus       50 VTGSFKEREAVRQLAKTCDVVTAEIEHVDTYALEEVASEVK---IEPSWQAIRTIQNKFNQKEHLRKYGIP 117
S.cerevisiae     49 VNGSFSNPLDIEKLAEKCDVLTIEIEHVDVPTLKNLQVKHPKLKIYPSPETIRLIQDKYIQKEHLIKNGIA 119
E.coli           45 -------------------VITAEIERWPETALTRELAR-H--PAFVNRDVFPIIADRLTQKQLFDKLHLP 93
V.cholerae       44 VLGHG-----LEQAIEQVDVITAEFEHIPHPILDLCARS-G--KLYPSAEAIKAGGDRRLEKALLDRAQVA 106
A.aeolicus       46 FRTGQ-----ISEFVDSCDIITYEFEHIKDEVLEKCES-----KLIPNPQALYVKKSRIREKLFLKKHGFP 106
M.leprae         58 VIGSHTDLEDLRRVALGADALTFDHEHVPTELLDKLVAEGI--NVAPSPQALVHAQDKLVMRRRLAALGAA 126
M.tuberculosis   62 VIGSHTDLAALRRVAAGADVLTFDHEHVPNELLEKLVADGV--NVAPSPQALVHAQDKLVMRQRLAAAGVA 130
B.subtilis       54 ITAHYNDREAIRKLAEISDIITYEFENIDYDALHWLKDH-A--YLPQGSELLLITQNRETEKKAIQSAGCE 121
S.aureus         54 IQAKYDDEKALNQLGQKCDVITYEFENISAQQLKLLCEK-Y--NIPQGYQAIQLLQDRLTEKETLKSAGTK 121
S.solfataricus   53 FSPQD-----YKEFVDSSDVITFEFEHVYEKALEYAEYS-G--KLLPRLNSVELKRERYKEKLFYRQHNLP 115
V.aconitifolia   49 MVGSFDESTKVEEFAKRCGVLTVEIEHVDVDTLEKLEKQGV--DCQPKASTVRIIQDKYQQKVALLPAWIP 117

B.anthracis     121 VATYRLVQN---QEQLTEAIAE----LSYPSVLKTTTGGYDGKGQVVLR-------SEADVDEARKLANAA 177
A.clavatus      118 MAEHRELVEN-TPAELAKVGEQ----LGYPLMLKSKTMAYDGRGNFRVNSQ-----DDIPEALEALK--DR 176
S.cerevisiae    120 VTQSVPVEQA-SETSLLNVGRD----LGFPFVLKSRTLAYDGRGNFVVKNK-----EMIPEALEVLK--DR 178
E.coli           94 TAPWQLLAE---RSEWPAVFDR----LGELAIVKRRTGGYDGRGQWRLRANETEQ---------LPAECYG 148
V.cholerae      107 NARYTMIRS---RDDLTSAIAE----IGLPMVLKSALGGYDGKGQWRLKEPTQIESVWQELAQYLAANPEQ 170
A.aeolicus      107 VPEFLVIK----RDEIIDALKS----FKLPVVIKAEKLGYDGKGQYRIKK------LEDANQVVKNHDKEE 163
M.leprae        127 MPRFMALDSVDDLAEIDAFAQRLTGSKDAPMVVKAVRGGYDGRGVQMVRDS-----AHAREVASGYLVDGM 192
M.tuberculosis  131 VPRYAGIK---DPDEIDVFAARV----DAPIVVKAVRGGYDGRGVRMARDV-----ADARDFARECLADGV 189
B.subtilis      122 VAPYSIVKT---KNELKQAVQE----LRLPAVLKTCRGGYDGKGQFVIK-------EEAQMEQAAALLEHG 178
S.aureus        122 VVPFISVKE---SKDIDKAIET----LGYPFIVKTRFGGYDGKGQVLIN-------NEKDLQEGIKLIETS 178
S.solfataricus  116 TPRFYVAED---GEEALKILREE---FNNVGVIKESKGGYDGKGQYFIFN------DVEKYQFLR--EKKE 172
V.aconitifolia  118 LPEFMKIDDL-----KAKKWDS----LDIHFMIKSRRLAYDGRGNFVAKSE-----EELSSAVDALGGFDR 174
```

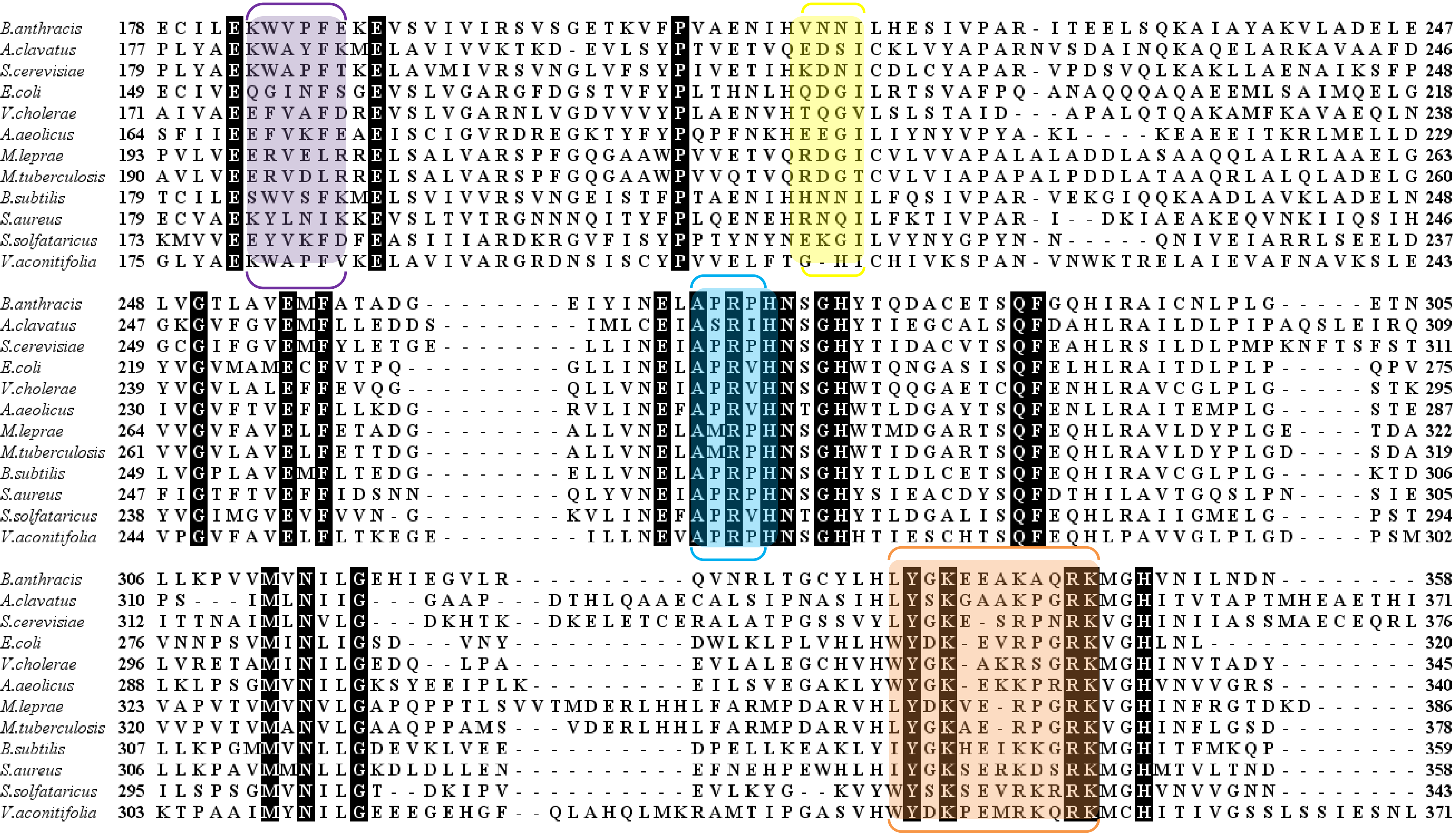
B.anthracis 178 ECILEKWVPFEKEVSVIVIRSVSGETKVFPVAENIHVNNILHESIVPAR-ITEELSQKAIAYAKVLADELE 247
A.clavatus 177 PLYAEKWAYFKMELAVIVVKTKD-EVLSYPTVETVQEDSICKLVYAPARNVSDAINQKAQELARKAVAAFD 246
S.cerevisiae 179 PLYAEKWAPFTKELAVMIVRSVNGLVFSYPIVETIHKDNICDLCYAPAR-VPDSVQLKAKLLAENAIKSFP 248
E.coli 149 ECIVEQGINFSGEVSLVGARGFDGSTVFYPLTHNLHQDGILRTSVAFPQ-ANAQQQAQAEEMLSAIMQELG 218
V.cholerae 171 AIVAEEFVAFDREVSLVGARNLVGDVVVYPLAENVHTQGVLSLSTAID---APALQTQAKAMFKAVAEQLN 238
A.aeolicus 164 SFIIEEFVKFEAEISCIGVRDREGKTYFYPQPFNKHEEGILIYNYVPYA-KL----KEAEEITKRLMELLD 229
M.leprae 193 PVLVEERVELRRELSALVARSPFGQGAAWPVVETVQRDGICVLVVAPALALADDLASAAQQLALRLAAELG 263
M.tuberculosis 190 AVLVEERVDLRRELSALVARSPFGQGAAWPVVQTVQRDGTCVLVIAPAPALPDDLATAAQRLALQLADELG 260
B.subtilis 179 TCILESWVSFKMELSVIVVRSVNGEISTFPTAENIHHNNILFQSIVPAR-VEKGIQQKAADLAVKLADELN 248
S.aureus 179 ECVAEKYLNIKKEVSLTVTRGNNNQITYFPLQENEHRNQILFKTIVPAR-I--DKIAEAKEQVNKIIQSIH 246
S.solfataricus 173 KMVVEEYVKFDFEASIIIARDKRGVFISYPPTYNYNEKGILVYNYGPYN-N-----QNIVEIARRLSEELD 237
V.aconitifolia 175 GLYAEKWAPFVKELAVIVARGRDNSISCYPVVELFTG-HICHIVKSPAN-VNWKTRELAIEVAFNAVKSLE 243
B.anthracis 248 LVGTLAVEMFATADG--------EIYINELAPRPHNSGHYTQDACETSQFGQHIRAICNLPLG-----ETN 305
A.clavatus 247 GKGVFGVEMFLLEDDS--------IMLCEIASRIHNSGHYTIEGCALSQFDAHLRAILDLPIPAQSLEIRQ 309
S.cerevisiae 249 GCGIFGVEMFYLETGE--------LLINEIAPRPHNSGHYTIDACVTSQFEAHLRSILDLPMPKNFTSFST 311
E.coli 219 YVGVMAMECFVTPQ---------GLLINELAPRVHNSGHWTQNGASISQFELHLRAITDLPLP-----QPV 275
V.cholerae 239 YVGVLALEFFEVQG---------QLLVNEIAPRVHNSGHWTQQGAETCQFENHLRAVCGLPLG-----STK 295
A.aeolicus 230 IVGVFTVEFFLLKDG--------RVLINEFAPRVHNTGHWTLDGAYTSQFENLLRAITEMPLG-----STE 287
M.leprae 264 VVGVFAVELFETADG--------ALLVNELAMRPHNSGHWTMDGARTSQFEQHLRAVLDYPLGE----TDA 322
M.tuberculosis 261 VVGVLAVELFETTDG--------ALLVNELAMRPHNSGHWTIDGARTSQFEQHLRAVLDYPLGD----SDA 319
B.subtilis 249 LVGPLAVEMFLTEDG--------ELLVNELAPRPHNSGHYTLDLCETSQFEQHIRAVCGLPLG-----KTD 306
S.aureus 247 FIGTFTVEFFIDSNN--------QLYVNEIAPRPHNSGHYSIEACDYSQFDTHILAVTGQSLPN----SIE 305
S.solfataricus 238 YVGIMGVEVFVVN-G--------KVLINEFAPRVHNTGHYTLDGALISQFEQHLRAIIGMELG-----PST 294
V.aconitifolia 244 VPGVFAVELFLTKEGE--------ILLNEVAPRPHNSGHHTIESCHTSQFEQHLPAVVGLPLGD----PSM 302
B.anthracis 306 LLKPVVMVNILGEHIEGVLR----------QVNRLTGCYLHLYGKEEAKAQRKMGHVNILNDN-------- 358
A.clavatus 310 PS---IMLNIIG---GAAP---DTHLQAAECALSIPNASIHLYSKGAAKPGRKMGHITVTAPTMHEAETHI 371
S.cerevisiae 312 ITTNAIMLNVLG---DKHTK--DKELETCERALATPGSSVYLYGKE-SRPNRKVGHINIIASSMAECEQRL 376
E.coli 276 VNNPSVMINLIGSD---VNY----------DWLKLPLVHLHWYDK-EVRPGRKVGHLNL------------ 320
V.cholerae 296 LVRETAMINILGEDQ--LPA----------EVLALEGCHVHWYGK-AKRSGRKMGHINVTADY-------- 345
A.aeolicus 288 LKLPSGMVNILGKSYEEIPLK---------EILSVEGAKLYWYGK-EKKPRRKVGHVNVVGRS-------- 340
M.leprae 323 VAPVTVMVNVLGAPQPPTLSVVTMDERLHHLFARMPDARVHLYDKVE-RPGRKVGHINFRGTDKD------ 386
M.tuberculosis 320 VVPVTVMANVLGAAQPPAMS---VDERLHHLFARMPDARVHLYGKAE-RPGRKVGHINFLGSD-------- 378
B.subtilis 307 LLKPGMMVNLLGDEVKLVEE----------DPELLKEAKLYIYGKHEIKKGRKMGHITFMKQP------- 359
S.aureus 306 LLKPAVMMNLLGKDLDLLEN----------EFNEHPEWHLHIYGKSERKDSRKMGHMTVLTND-------- 358
S.solfataricus 295 ILSPSGMVNILGT--DKIPV----------EVLKYG--KVYWYSKSEVRKRRKMGHVNVVGNN-------- 343
V.aconitifolia 303 KTPAAIMYNILGEEEGEHGF--QLAHQLMKRAMTIPGASVHWYDKPEMRKQRKMCHITIVGSSLSSIESNL 371

Figure S2:

Shown is a series of omit maps that were generated by setting the occupancy of the bicarbonate atoms to zero and re-refining the structure **3V4S-B**. The top panel shows all three maps; mF_o-DF_c (dark blue), $2mF_o$-DF_c (orange) and $3mF_o$-$2DF_c$ (teal) at 3.5, 1.5 and 2.5 σ, (0.28, 0.43, and 0.78 $e^-/Å^3$) respectively, from a top-down view. The enzyme (green) is depicted as a cartoon representation, with the side chain of R272 shown in a stick representation as an anchor for orientation. The bottom three panels show the individual maps; mF_o-DF_c, $2mF_o$-DF_c and $3mF_o$-$2DF_c$ (left to right) rotated 90° to give the side view, confirming the correct trigonal planar density.

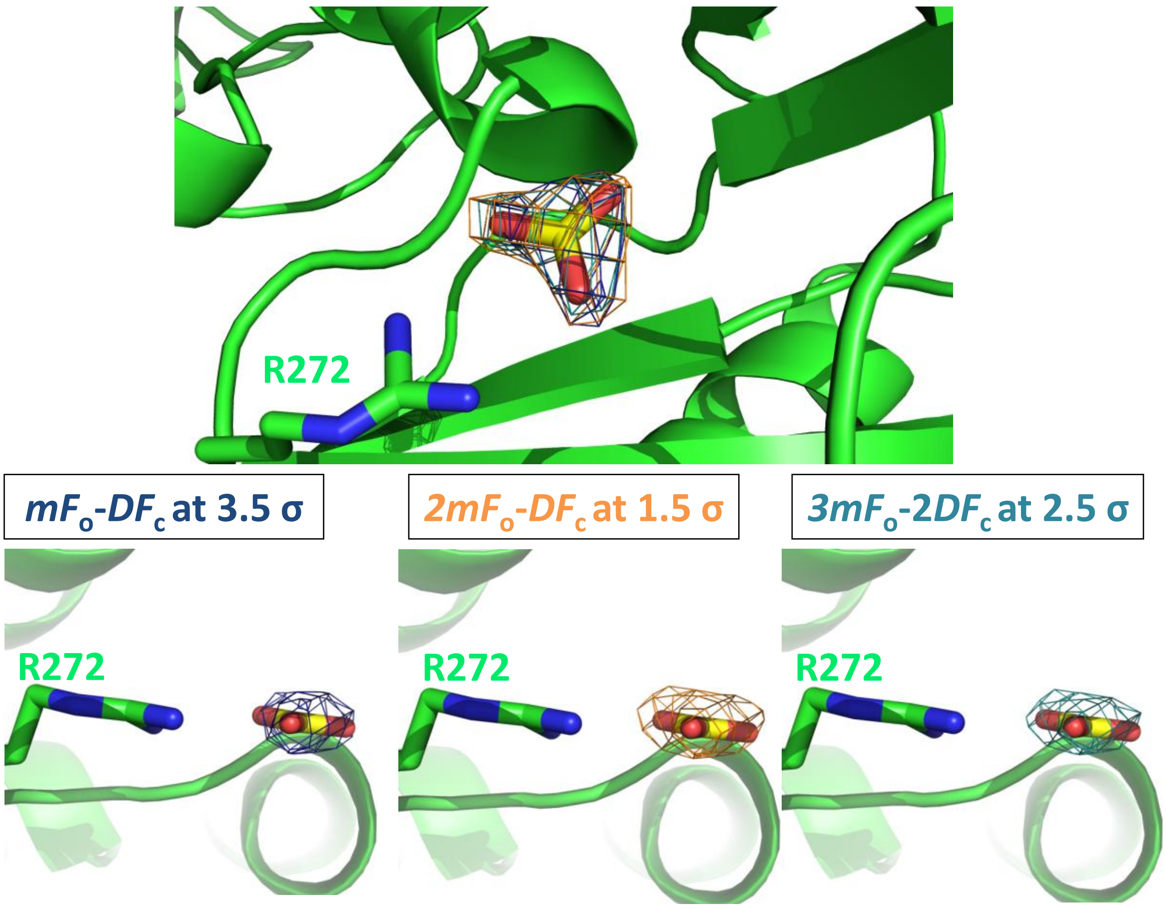

Figure S3:

Shown is a mF_o-DF_c omit map (dark blue mesh) focused on HCO_3^- (red and yellow) and its proximity to R272 (cyan) and K348 (orange) from **3V4S-B**. The map was generated by setting the occupancy of all the atoms from HCO_3^-, R272 and K348 to zero and re-refining the structure. The mF_o-DF_c map is contoured to 3.0 σ (0.24 e^-/Å^3).

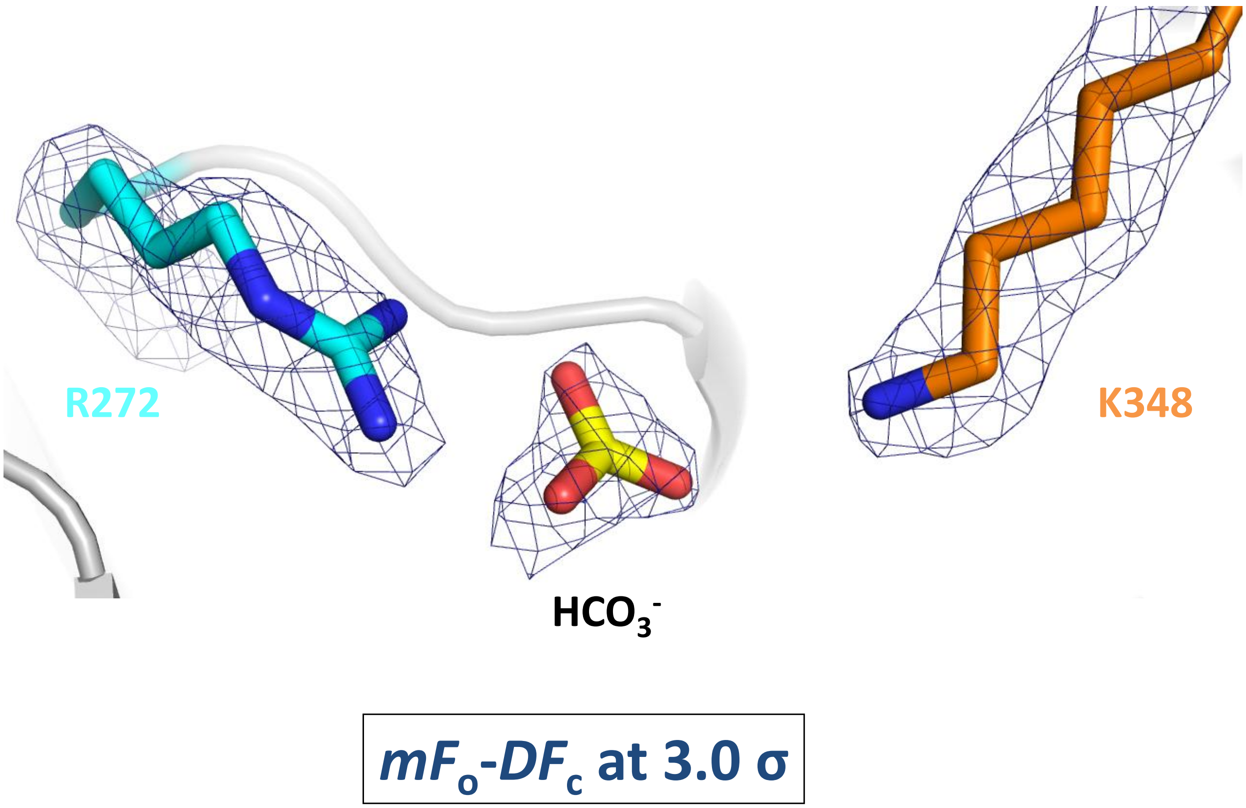

Figure S4:

Shown is the environment of water 479 from **3R5H-B** (same position as water 41 in **3V4S-B**) that acts as a base in our mechanism, we propose that a hydrogen is removed by K348 (orange) to allow water 479 to deprotonate the exocylic amine of AIR. The water displays typical tetrahedral geometry, interacting with the O2 atom of bicarbonate and both of the amine groups from AIR and K348 of the modeled active site. It also hydrogen bonds with water 385 of **3R5H-B.** Dashed lines are possible hydrogen bonds available within 3.4 Å.