



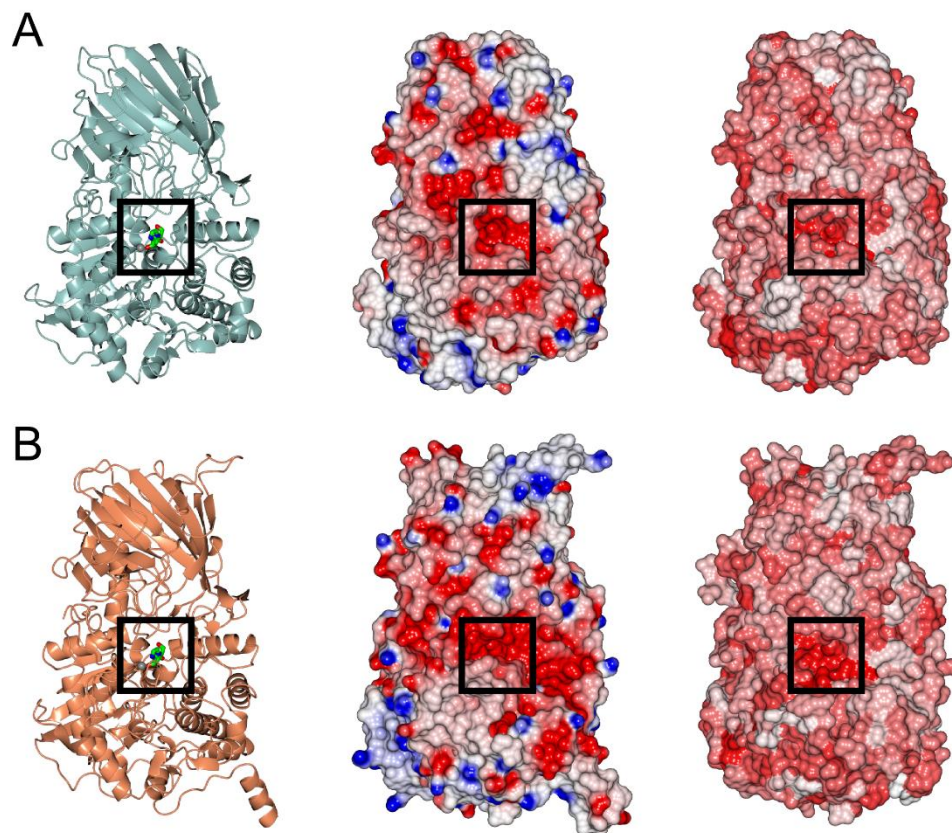
STRUCTURAL  
BIOLOGY

**Volume 74 (2018)**

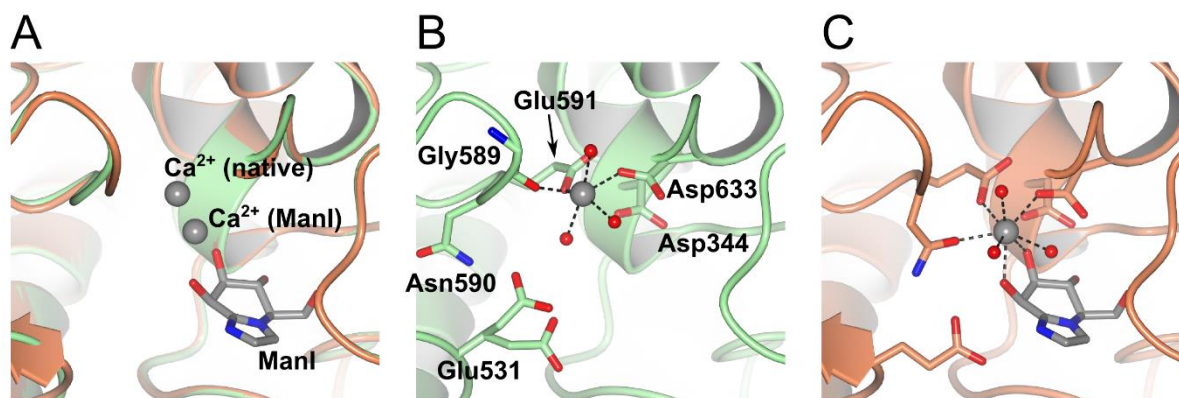
**Supporting information for article:**

***Bacteroides thetaiotaomicron* generates diverse  $\alpha$ -mannosidase activities through subtle evolution of a distal substrate-binding motif**

**Andrew J. Thompson, Richard J. Spears, Yanping Zhu, Michael D. L. Suits, Spencer J. Williams, Harry J. Gilbert and Gideon J. Davies**



**Figure S1** Alternate structural representations of BT3130 (A, green ribbons) and BT3965 (B, orange ribbons). Panels, from left to right, show ribbon projections of respective enzyme-ManI complexes (ManI and a bound  $\text{Ca}^{2+}$  ion are drawn as enlarged cylinders and spheres respectively), enzyme-ManI contoured as an electrostatic surface, and finally native enzyme structures coloured according to sequence conservation using the expanded alignment shown below in Figure S3 (areas of strong conservation are indicated by intense red shading). The location of the active site pocket is indicated by a dark box in each model. While amino acids in both molecules are generally well conserved, strongest similarity is observed within the inner core of the substrate binding pocket. Sequence alignments were mapped to coordinate files using the Consurf server (Ashkenazy *et al.*, 2016, Landau *et al.*, 2005).



**Figure S2** Overlay (A) of  $\text{Ca}^{2+}$  bound in the active site of BT3965-native (light green, and panel B) and BT3965-ManI (orange, and panel C).  $\text{Ca}^{2+}$  is aberrantly coordinated in native BT3965 (B), and although both clear octahedral coordination geometry and an observed strong difference map peak in the absence of a modelled ion are suggestive of  $\text{Ca}^{2+}$  binding, *B*-factor analysis compared to surrounding atoms suggests only approximately 50-60% occupancy at this site. Displacement of  $\text{Ca}^{2+}$  away from the ideal position for coordination and catalysis yields a dual, unnatural conformer in the catalytic acid, Glu531. Supplementation with exogenous divalent salts and binding to ManI (C) brings about a subtle conformational change, rotating Asn590 into the active site pocket, positioning both  $\text{Ca}^{2+}$  and Glu531 into the activated state.

BT1032 ---MRR-----ILLTCTLAFTLLPVFGGEGKAPMDKEKSLLDYVDPFIGTT-----  
 BT1769 -----MKLTHILASAAIISTLSA--CNG-----SLQTADRTPDVYVNPYIGNISH-----  
 BT1878 MIIMKRKSPVMV---ALAGLFMYTS--CSP---TEQAEKTKDYQVYNTFIGAA-----  
 BT2199 ---MVTR---LFFLSIC--FPFLTS--CQ-----QSKKTEFVYVNLMTGTESTFA-----  
 BT2629 ---MKTHFSFKHLLFLG--GAVL--Y--SL-----QSSAVKNPDVYVSLTGTQSKFE-----  
 BT3130 -----MKK---LILLSVCAFSL--TLA-----TLQAGEITRYVNPFIGTGAIDG-----  
 BT3527 -----MKKLALLA--FTLFSAM--NM-----DAKTIITGPDVYVPLVGTQSKHA-----  
 BT3773 ---MNAFKL---LMAASCFGL--ATSGVAKQDYKSEGLLQYVDPYIGSG-----  
 BT3784 ---MKTHFSFKHLLFLG--GAVL--Y--SM-----QISSAVKNPDVYVSLTGTQSKFE-----  
 BT3858 ---MMNRRLNKKRVGSCFLMAMAFS--CTH---TDQYPTKDPVYVNPYIGNISH-----  
 BT3962 ---MNRKLLSIT---MILLG--VS-----SIFAQKQPDVYVNLMTGTESTFA-----  
 BT3963 ---MNRKMK---KVLMTM--AAV---TL--SVAQAQPDVYVNPYIGNISH-----  
 BT3965 -----MAQTEKLTQVYVNPFIGT-----  
 BT3990 ---MKRLLIA---CILNCTL--LS-----QAKDWTQVYVNPYIGNISH-----  
 BT3991 ---MKLKTLLIGCGPGFIMLNS--CTE-----SPSVKDYSAVNPYIGNISH-----  
 BT3994 ---MKTPIY---LLLIVCIFAS--CNT---KQTAEDYTSVYVNPYIGNISH-----  
 BT4073 ---MYKASANILACSFLWLTAA--CSSA---GVVTEDLIPTDYVNPYIGNISH-----  
 BT4092 ---MKKLVYLI---IFFLY--S--TV---MNAQLKDLVQVYVNPYIGNISH-----  
 BT4093 ---MK-----HK--FLFLLFSLVLEGM--V-----TTQAVAGDYVHVNTLIGTKGTG-----

BT1032 ---NFGTTNPGAVCPNGHMSVFPFNVMG--SADNTYDKDARWWSPTYEHTNCFPTGYSH  
 BT1769 ---LIVFPTFLQLPNSMLRYPERAD-----YSELKGLPL  
 BT1878 ---DNHFTPGACYFPGMIQTSPTV--GAV-----GWRVCSEYTYQDSLWGFPTQ  
 BT2199 ---FSGHNTYPAVAVPWGMNFWSPTQGEN-----GSGW---MYTYDLSMRGRFQ  
 BT2629 ---LSTGNTYPATALPWGMNFWPTQTKM-----GDGW---AYTYDADKIRGFQ  
 BT3130 ---GLSGNNYPGATSPFGMIQLSPDTSEAP-----NWGDASGYDNRNPTIFGFSH  
 BT3527 ---LSTGNTYPALALPWGMNFWPTQTKM-----GDGW---AYTYDADKIRGFQ  
 BT3773 ---FHGHVFGTSPVYGMVQLGPNN--IHK-----GWDWCSGYHSDSLILGFSH  
 BT3784 ---LSTGNTYPATALPWGMNFWPTQTKM-----GDGW---AYTYDADKIRGFQ  
 BT3858 ---LLVPTYPTVHLNPSMLRVPYPERGD-----YTSRNVGLPV  
 BT3962 ---LSNGNTYPALALPWGMNFWPTQTKM-----GDGW---AYTYDADKIRGFQ  
 BT3963 ---GMHFTPGACTPFGVQLSPDTDTI--PHNVNGAYQKNAYEYCAQYRDKTIVGFSH  
 BT3965 ---GYGNVYGAQIPFGGIQISPDTSR-----FYDAASGYKNHLLTMGFSL  
 BT3990 ---LSTGNTYPALARPWGMNFWPTQTKM-----GDGW---QYTYTANKIRGFQ  
 BT3991 ---GHHTFPGAQVPHGMIPSPDT--RID-----GWDACSQYDNDTNGFSH  
 BT3994 ---LSTGNTYPALALPWGMNFWPTQTKM-----GDGW---QYTYTANKIRGFQ  
 BT4073 GYVHGLKTEPGATTPYGMVQVSPNIT---GLP-----GWDRIISGYFYDSTLIGFSH  
 BT4092 ---LSVGNTPYPTGMPYGMHMSAQTKGN-----GDGF---RYMVAVDKIRAFSQ  
 BT4093 ---LTSGLYYPGATYPMVQVTPSPYS-----KRSQFVI

BT1032 VNLGSGVC--PELGSLLMPTTGELNVD-----YKEYGSKYK--DEQASPGYSNYLTK  
 BT1769 IVTNHRE---RSAPNFSYQEKLR-----PVIYTNNDNEHITPYSPDVELDD  
 BT1878 THLNGTGC--MDLGDILVMPVGTTRAR--A---WD---AYRSHFPKDKAATPGYTVLSD  
 BT2199 THQSPWMI--NDYQFSPIMPLAGEKMS-----HKRLVFP--SHQEKATYNYVSTFN--  
 BT2629 THQSPWM--NDYQFAIMPITGLVFD-----QDRRASWF--SHKAEVAKPYVYKVLAD  
 BT3130 THLSGTGA--SDLDITLMPSTSGR-----TSSAFTDHEEKARPGYVQVMD  
 BT3527 THQSPWMI--NDYQFSPIMPLAGEKMS-----QDQRASWF--SHKAEVAKPYVYKVLAD  
 BT3773 THLSGTGC--FDLGDILVMPVGTTRAR--RGMQDDIRDGYASKYSHDNEIARYPYSLDR  
 BT3784 THQSPWM--NDYQFSPIMPLAGEKMS-----QDQRASWF--SHKAEVAKPYVYKVLAD  
 BT3858 VHTSHRG---SSAFNLSVQGEVSR-----PIVSYDYLENITPYGYSVLDE  
 BT3962 THQSPWMI--NDYQFSPIMPTKQLKID-----QDSRASWF--SHKAEVAKPYVYKVLAD  
 BT3963 THLSGTGH--SDLDILLMPVAGDVKLNPGRADHPEEGYRSRFDHATEKATPGYVEMLDD  
 BT3965 THLSGTGI--PDLGDFLIPGTEGMKLEPGTHEPDDQGYRSRFDHATEKATPGYVEMLDD  
 BT3990 THQSPWMI--NDYQFSPIMPLAGEKMS-----EKRASWF--SHKAEVAKPYVYKVLAD  
 BT3991 THVSGTGC--CDYGDVLLMPTVQKPYLTDPESSQKLAAYASAF--SHENETAEPGYVSLFD  
 BT3994 THLSGTGA--GDLYDISPMPVTLVYKE--A---EAPLGIYSKFSHDEESAYAGYVQVRLD  
 BT4073 TOMSGVWGFGLDGLNVLMPPTGELQKLAGKEDGSIKGYRSDYKATETAKAGYVSLVLD  
 BT4092 SHQCSPWV--SDYAVYSFMPMVGELVVN-----QDARATKFSHDEIAKPHYKVTFD--  
 BT4093 NQLSGGGC--EHMGNFPTFPVKGKLMSPDN---I---LNYRINVSSEKGHAGYEMVQ--

BT1032 YNITEVSATPRTGIARFTFPKGES--HILLNLEGLTNE-----GAMLRV--S  
 BT1769 NRMKAEYALSHQSAIYRITYEADKPAYLVNRSNG-----SIHAN  
 BT1878 PQVKAEKTASIIHAALHRYTYHKADASALLIDLQHG---PAMREEQYHSQVNSVEVNE--D  
 BT2199 NGLQTSLSATSRGAVFVSPPEKEDQYVVVDAYNG---G-----SITITIEPE  
 BT2629 HDVTEELAPTERAVFRFTYPTKNAVYVDAFDK---G-----SYVKVIEPE  
 BT3130 ENINAELTTQRNGIHRVYQYAGKDAEII--LDMHSDADGSGWRRIINSQI-----RIL--N  
 BT3527 HDVTEELAPTERAAAFRTFPPEHDSYVVVDVAFDN---G-----SPVKVIPS  
 BT3773 YNKAELTATDRVGFHRYTPEGKPVSI--LDLREGNSNAY---DSYI-----RNV--D  
 BT3784 HDVTEELVPTERAAMFRFTYPTKNAVYVDAFDK---G-----SYVKVIEPE  
 BT3858 ADIQEYAPSHOGLVHISFTEGDNALVNTKMG-----KVAE  
 BT3962 YNMTTELPTEKAYFRFTEADSAVYVDAFDK---G-----SYVKVIEPE  
 BT3963 YGKKAQATQRTGVHRYTPEGKDGHL--LDLVHGIY---NYDGVLANLWLVN--N  
 BT3965 YGVKAEMTSGRVSRGFRFTYPTESDNAPIMIDMNHNLN-----Q---SCEWNLNMI--N  
 BT3990 HDIVTEMTPTERAVFRFTYPTPEHDSYVVVDVAFDN---G-----SYKIIPE  
 BT3991 YQVKAEIATKRGAIHRYTPESTESGPIIDLDSLQRQ---TNYE---MEIEVI--S  
 BT3994 YHINVELTATERCGIQRVYFPAKAAAL--FLNLKKA---MNWD--FTNDSHI---EVV--D  
 BT4073 YKIKVSSATPHCGILQFTFSPNEQSRIQIDLARRVGGTS-----TSOYKVL--D  
 BT4092 NGITTEMAPTRGVHLRFSPYTPGDAYLVLDGYTD---M-----SEIKIDPA  
 BT4093 EDIKAKLTVTERTGMASVEYTPADQYGTIIIG--GGISAT-----PIEQAA--IVITA

BT1032 DSEIEGMKLL---GTFYCY--NPQAVFPIYFVHRVNVKVPATGYYWKRQRPMTGVEAEWDRD  
 BT1769 ENFTSGRGLL---PFWV---DNVNVVIEEQRERF--SAGLEMTLET-----  
 BT1878 AQTLAGHNV---TVMV---DQYFVFMKFRVYVDSLLDPMQY---  
 BT2199 KRLVKGATRY---NN---GGVDPNANVFMMEF--SHVPIEVGTVNGD--TL-----  
 BT2629 ENKIIGYSTK---NS---GGVDPNANVFMVFIQDFKPTFVSTVFEN--NI-----  
 BT3130 DHAVEGRII---TGWAK---LRKIYVMEFSSPILSTLRDGGRV-----  
 BT3527 ENKIIGYSTK---NS---GGVDPNANVFMVFIQDFKPTFVSTVFEN--NI-----  
 BT3773 DYTVEGYRYV---RGWSP---SRKVFVFLKDKQIEKFTAYDNDTPKL-----  
 BT3784 ENKIIGYSTK---NS---GGVDPNANVFMVFIQDFKPTFVSTVFEN--NI-----  
 BT3858 EKGVSQYQVI---DN---TPTKIYLYLETSQLPLRKGVLADGKVDM-----  
 BT3962 ENKIIGYSTK---NS---GGVDPNANVFMVFIQDFKPTFVSTVFEN--NI-----  
 BT3963 DTLTGYRIT---NGWAR---TNYTYFAISLSQPKIDYKDK--EKVLYNGF--W--  
 BT3965 DSTITGYKLV---KGWGP---ERHVFYATFSSKLTGLRFVQDKKPVYNTS--R--  
 BT3990 ENKIIGYSTK---NS---GGVDPNANVFMVFIQDFKPTFVSTVFEN--NI-----  
 BT3991 DTEICGHKKT---TYWAF---DQYINFYARF--SKFPAFTLITDVSMT-----  
 BT3994 SVTIQGYRYS---DWAR---DQYIFRFRF--SKFDPKVELDTAII-----  
 BT4073 DYTICQWMCPTDQGGWNGECSYVYVYVQF--SKPLSNYGLVWGSADIPDEWVRK--RDE  
 BT4092 KRQISQWVNN---QR---FVNDKSFVNVVQF--KPFEDYGVWENKQDEV-----  
 BT4093 PNKCEGYAEG---GNF---CGLRTPYRVYVAFEDTDAFETGTWKRREL-----

BT1032 ---QGKLYTRYGKEIAGDD---VGAFTFETEEGEQVEMQGVSVSIVENARL  
 BT1769 -----SKDNAEGAN---ACAARWA--DGTFTVNLRYGISFISEEQAEK  
 BT1878 -----EKGKR---I--L--ATF---DLKPGDELMMKVALSTTSVEGAKK  
 BT2199 -----LHHQTDVAADY---TCAYLKPDPVFAKELTIRTAASSFISPEQAAI  
 BT2629 -----LPNETEAKGNH---TGAVIGFATKKEIVHARVASSFISPEQAE  
 BT3130 -----HENTAINGTN---LHGCFFGQLNGKPLTCKVALSVSMENARE  
 BT3527 -----DTPNKLEANDKH---AGALIGFKTRRGEQVNVKVASSFISPEQAE  
 BT3773 -----WQLKV---ES---VRSVLTFF---GWEKVKIKVAISSVSDNAMH  
 BT3784 -----LPNETEAKGNH---TGAVIGFATKKEIVHARVASSFISPEQAE  
 BT3858 -----ESKEG---SALALY--GSEKNLRYGIFISABQAKK  
 BT3962 -----VETHELEQSNH---VGAAGIFSTKKEGQVHARVASSFISPEQAE  
 BT3963 -----RRFKL--EKNFPEITGRK---IVAVNFETAKDELVVKVALSAVTEGAVK  
 BT3965 -----FR-----SSVEAWGN---LMACISFDTKAGEVTVKTAISAVSDGARN  
 BT3990 -----QENVAEQTDDH---AGALIGFKTRRGEQVNVKVASSFISPEQAA  
 BT3991 -----DNKRLPVCKAVLHFNKDKDEVLVKGVSADVIAAGARK  
 BT3994 -----KDKQ--HIGTA---VIARDFHTEEGEQLVNTAISVSMEGAARK  
 BT4073 VSIPLYLTRISQAPV--IKDKKELEKGH---LGFTEFPTKEGEQVEMKVISFVMEGAAN  
 BT4092 ---FPQKLDGAGK---YGAFIGF---KGSXVQAKAASSYISAEQALI  
 BT4093 ---MPNTTFAEGEY---SGVYFTFDVNNKKNIQYKIGSVYVSENARE

BT1032 NLDKEOSGK--NFEQVLSARAQWDDLSRITVEGGTDAQKTVFYALYHLLIHPNVLQDV  
 BT1769 NLHRELKDY--NIKALAEAGRIWNETLGRIQVEGGTDDKTVFYSFYRTERPFCMSEG  
 BT1878 NLQAEIPDW--NFDQVLAADHDWNSVLSRMDVEGTDDE--RTNFYTCFYHALIQPNQISDV  
 BT2199 NFNREVADA--DVQLISGKAREQWNNYLRVVEAGGTDQELRTFYSLYRTRLLFPREFYEF  
 BT2629 NLKE--LGNK--SFDQVLANGREIWNREMSKIEIEDDNIOLNRTFYSLYRTRLLFPREFYEF  
 BT3130 NMEQEAHPW--DFDRYVAADADWEKQLGKIEVKGTEVO--KEIFYALYHMTIQPNMSDV  
 BT3527 NLKE--LGKD--NIDQIAAKGRKVVNDVLRGIEVEDDDIHLRTFYSLYRTRLLFPREFYEF  
 BT3773 NLQAELSHW--DFDKVVKMSDRWVKQDKMTEVESDDEAAKRVFYTAHYHTMIAPTYLQDV  
 BT3784 NLKE--LGKD--SFDQVLTGKGDWNRMSKVVEDDNIOLNRTFYSLYRTRLLFPREFYEF  
 BT3858 NLQRDITTY--DVKAADAGRRIWNTLKGKIEVEGSEDEKEIFYSLYRTERMINLSED  
 BT3962 NLKE--LGNK--TFEQTKAAGKAWNDVLRGIEVEDDENRMRFTFYSLYRTRLLFPREFYEF  
 BT3963 NLRAEASGK--SFEGLAEARTDWNNELDHFEIEGTPDQ--KAMFYTSLYHTMNPVSYMDV  
 BT3965 NKW--ELDCL--TWEELKAGCALWEKLEKVTIADRK--KEPFTYSABHAALHPFIQDS  
 BT3990 NNNE--LGRD--NIEQLAQKGDAMWVLRGIEVEGGLDQVRYFTFYSLYRTRLLFPREFYEF  
 BT3991 NVESEIPEW--DFDKVRKDARTANNYLSKIDITTSKEDKTFYALYHTAISPNLFTDA  
 BT3994 NLQAEVPEW--DFDKYLAETKANWNRQLGKIEVEGDNQDDKVNFTYALYHMTIAPTYSDV  
 BT4073 NFKQEIASK--NFAQVKQASDLWNKELGRISRISGGTDEKTVFYSLYHTMIDPRIYTDV  
 BT4092 TLNKLKGLKDNLEVTKARGQKTWNEVLRNIVVEGTEQDKMFTFYSLYRTRLLFPREFYEF  
 BT4093 NLKAENAEW--DFQIKQNAEAKWNNYLGKIEVEGTPDRTPQFTYHLYRSFIHPNVCSDV  
 .. \* : . : \* : . :

BT1032 NGEYPA----MESDQIL--TTGTRYTVFSLWDTYRNVHQLLTLVYPERQMEMVRTMLDM  
 BT1769 ----GRYSAFDGKVEH--DNGTFPYTDWINDTYRAAHPRLTLIDQKKEEDIASYLDM  
 BT1878 DGMYR----AADSIVK--AGTGFYTFSLWDTYRAAHPRTYTLVYPERVDFVNSLIEQ  
 BT2199 DQGNVYTFYSPDQVH--DQMTWQVDFTRAVHPLTLLYVESESRVPSIINA  
 BT2629 DAKQVHYSPYNGEVR----PGYMTDTGFWDTRCLFPFLNLMYSPMNMKQEGLVNT  
 BT3130 NGEYMA----ADYTRKVANNEYHTYFSLWDTFRASHPLYTLLEPERVDFVNSLIRQ  
 BT3527 DAKGDIHYSPYNGEVL----PGYMTDTGFWDTRCLFPFLNLMYSPMNMKQEGLVNT  
 BT3773 NGEYRG----MNDMIYT--DPKKANYTTLWDTYRALNPLMTIIQPEMNDVNVNSMLSI  
 BT3784 DAKGQVHYSPYNGKVL----PGYMTDTGFWDTRCLFPFLNLMYSPMNMKQEGLVNA  
 BT3858 ----GKYSAFDGKIEH--DGGVFPYTDWINDTYRATHPLRLILIEPQKELMIRSYIRM  
 BT3962 NAKGETVHYSPYNGEIR----PGYMTDTGFWDTRCLFPFLNLMYSPMNMKQEGLVNT  
 BT3963 DGSYRG----LDNHIHQ--AKGFTNYTIFSLWDTYRAEHPFLNLMYSPMNMKQEGLVNT  
 BT3965 DGOFRG----LDKNIK--AEGFTNYTIFSLWDTYRALHPFLNLMYSPMNMKQEGLVNT  
 BT3990 DANGQPIHYSPYNGQVL----PGYMTDTGFWDTRCLFPFLNLMYSPMNMKQEGLVNT  
 BT3991 DGRYLG----MDLEVHGDDTINPLYTVFSLWDTYRALHPLMTIIDPMLNNOFINSLIRK  
 BT3994 DGAYIG----PDKVHQ--SDQWVNYTTFSLWDTYRAAHPRTYTLVYPERVDFVNSLIEQ  
 BT4073 DQKIFIG--GDKVHQ--SDQWVNYTTFSLWDTYRAAHPRTYTLVYPERVDFVNSLIEQ  
 BT4092 KENGEFYHYSPYNGKIV--DGVMYTDWINDTYRTHPLNLMYSPMNMKQEGLVNT  
 BT4093 NGEYMG----ADRFVHK--SRSKHYTSFNWDTYRQIQLLSMLDPEVADVISHLQF  
 : \* : . : \* : . :

BT1032 YRE--HGWFPKWELYGRETLMEGDPSPVIVDTWMMGLRDFVDLAYEAMYSATLP--  
 BT1769 AEQMGMMMPFPEVTGTRRMSNHAVATVADALAKGLK--VDTAKAYEACRKGIEEKT  
 BT1878 GEV--QCFLP--IWLWKGNEFCMIGNHGVSIIEAYRKGFRGDAERAFNMVKTQTVS--  
 BT2199 YNE--SGFPEWASPG--HRGCMIGNNSVSLVDAWMMKGIQTVDAEKALEAMHQTQAR--  
 BT2629 YKE--SGFLPEWASPG--HRDCMVGNNSASVADAYIKGLRGYDIETLWEALKHGANA--  
 BT3130 YEY--YGFLP--IWLWQGDNYMCI GNHSPVITDAILKGI PGIDMEKAYEVNSVTS--  
 BT3527 YRE--SGFLPEWASPG--HRDCMVGNNSASVADAYIKGLRGYDIETLWEALKHGANA--  
 BT3773 YRQ--QDKLP--IWLWQGDNYMCI GNHSPVITDAILKGI PGIDMEKAYEVNSVTS--  
 BT3784 YLE--SGFLPEWASPG--HRDCMVGNNSASVADAYIKGLRGYDIETLWEALKHGANA--  
 BT3858 AEQSDRRMMPFPEVTGDSHRMNGNHAVAVINDAYCKGLKDPDLAAEYACKGALTEKTL  
 BT3962 YLE--SGFFPEWASPG--HRGCMVGNNSASVADAYIKGLRGYDIETLWEALKHGANA--  
 BT3963 EQQSVHGMPLIWSLMGNENCMGSHAVSLADAITKGVFSN--VDEALSAMVSTSTVP--  
 BT3965 YDKSVEKMLPISWFSYNETWCMIGYHAVSLADAITKGVFSN--VDEALSAMVSTSTVP--  
 BT3990 YLE--SGFFPEWASPG--HRGCMVGNNSASVADAYIKGLRGYDIETLWEALKHGANA--  
 BT3991 HQE--GGIYPMWDLASNYGTMI GYHAVPVIDAYMKGRNDAEAYKASLRAEYD--  
 BT3994 FEQ--NGRPLVWNFYGSSETDMI GYHAVPVIDAYMKGRNDAEAYKASLRAEYD--  
 BT4073 ADQSRREYERWELLNSYSGCMIGNPALSVLADAYMKGIRTYDVEKAYAVNTSAK--  
 BT4092 QEQ--CGWLP--SWSVSG--ETGMLGNHSTSLADAWAKGIRDPEKALKAYAEAMNK--  
 BT4093 AEQ--SGGSFPRWMMANIETGVMOGDPPTILIANAYAFGARNYDKPKIFKMRKGAEEP--  
 : \* : . : \* : . :

BT1032 GAE-----NLMRPDNDYMSKGYVPLREQ-----Y-----DNSVSHALEYYIA  
 BT1769 APWSGAPAGWL-----DNFYRENGYIPALRVDEPDPNVHPEKRPVAVTLGTSYD  
 BT1878 HFL-----KSDWEVYTKYGYFPDGLT-----X-----AESVSSLTLESYVD  
 BT2199 HALEIA-----SVGRDGFYDKLGIYFPF-----Y-----PEADAKLEIYAYA  
 BT2629 HLRGT-----ASGRLEYSYHOLGYVANNIG-----I-----GQVARTLEYAYN  
 BT3130 HPN-----SFFWEKYGFMPEINQ-----I-----TQSVSLTEQAFD  
 BT3527 HPQVS-----STGRLEGYDYNKLGYPVYVNG-----I-----NENAARTLEYAYN  
 BT3773 GQN-----G--VPYVMKAGYIPADKI-----I-----HEATSIAEMEYAVD  
 BT3784 HLRGT-----ASGRLAYDAYNKLGYPVANNIG-----I-----GQVARTLEYAYN  
 BT3858 LPWLRCPTEL-----DKFYQKQFPALNPGEEETKAVHSFERRQAVAVMLGNCYD  
 BT3962 HPRVS-----TGRRGEYYNKLGYPVYVNG-----I-----NENAARTLEYAYD  
 BT3963 YY-----EGVADYKMLGYIPLDKS-----I-----GTAASSTLEYAYD  
 BT3965 NY-----DCLPEYREMGYVFPDKE-----I-----AESVSKTLEYAYD  
 BT3990 HPEVS-----STGRLEGYDYNKLGYPVYVNG-----I-----NENAARTLEYAYD  
 BT3991 TTGIRCPDLVPLHMLPKAKYKNAIGYPCDRE-----I-----NESVAKALEYAYD  
 BT3994 DNY-----RGILYKOLGIYFPFWT-----I-----DHNENBSLSKTEYAFD  
 BT4073 GN-----DGLCYTPE-----I-----PLSEYTEYAYA  
 BT4092 GWGCG-----ANRGRFWEYFOLGYVYVNG-----M-----GS--SAQTHEYAYD  
 BT4093 GSKSQ-----DVETRPLKQVLDKGYNN-----I-----ASIQLEYTSA  
 : \* : . : \* : . :

BT1032 DFALSRFADALGKKKDA--EMFYKRSGLYKHYKSEFGTFRPILPDGTFYSPFNPROGEN--  
 BT1769 QWCLSRIFAALNKKEEA--EYLKCSYNYRNLNFKETAFFHPKDKQEGWIEPDPYFPGGM  
 BT1878 DYAAADAMRMGKKEEDA--AYFAKRADYKLNFDQSTNFMRPKADGTWKSFPNPSDVGH  
 BT2199 DMCILARFAESLQKQDIA--DQYQKAPNYRNLNYPHFGFMWTKDAKGNWRDRDATEWGG--  
 BT2629 DMWITLGLKGLKPENEIDYKXKHALNYKHYVPERKLMVGNKDNKGNVFNPDVADWGS--  
 BT3130 DMCVQQLAALNKDADY--QRPHKRSEYRNLHFPKTKFQSKNDKGEWIEPDPYQYGG--  
 BT3527 DMC1YQLGKALNRPKKEIEIFAKRAMNYKLYDPEHKLMRKNEDKGFQSPFNPLKWDG--  
 BT3773 DMGIAAMQKMKKADY--EYFKRAMNYKLYDPEHKLMRKNEDKGFQSPFNPLKWDG--  
 BT3784 DMWITLGLKGLKPENEIDYKXKHALNYKHYVPERKLMVGNKDNKGNVFNPDVADWGS--  
 BT3858 NMLAQIARTLAKDDY--KFMMSYTYRNVYNAETGFFPKNDKGFIEPDPYRYSGGQ  
 BT3962 DMC1YRMKGLRPAEELYSRQSNYRNLNFDPEKLMRKNEDKGFQSPFNPLKWDG--  
 BT3963 DMTIYQTLAKGNKEIA--DTRKRALNYRNIYDTSIGFARPRYSDGTFKKEFDVLOTYGE  
 BT3965 DYCIAQAAKLKGKEDDY--HYFLNRALSQYTLIDPETKMYRGRDSEKGNWRTFPTVAYQGP  
 BT3990 DMC1YRLAKELKRPKKEISLFAKRAMNYKLNFDKESKLMRKNEDKGFQSPFNPLKWDG--  
 BT3991 DMC1SIFAEAMNDFENK--AKYERFAKAYEFYDKSTRFMRGSDNLSGNWRTFPTVAYQGP  
 BT3994 DYCIAEMANKMGKKEIA--DEFYKRSQNYKVNYPATSFQMRDKNKGFIEPDKFAEYTP--  
 BT4073 DMCVQQLAALNKKKEEDA--KRFYKQAYRNMFDKAEKGFWRPNADGSKWAKPENAAAA--  
 BT4092 DFCGYQLAKMVGKHYQ--EVFARQMYNYKVNFDKSI GFMRGKVDGKQWEPDPLEWGG--  
 BT4093 DFAIQGFALHVGDEFASWRYFHARSWKNLNYSDTGWLQSRNPDGSKSL-----  
 : : . : . : . :

BT1032 --FEPNPGFHEGNSWNYTFYVPHDYGVLAKLGGKFPVNLQMVDEGL-----  
 BT1769 GA---REYEGNNGWYRNDVPHNADILSLMGNEQFIANLDRTPTEPLGRSKY-----A  
 BT1878 --AESTGGDYTEGNAWQYTHWVQHPQGLIALFGGEEFPLNKLDSLPTVKL-----  
 BT2199 -----PFTEGSSWHWTSVFDHPEGLSELMGGHEPMIARLDSMFVAPNTYNGYGFV  
 BT2629 -----EFCENSWHWSFCVFDHPQGLIDMLMGKKEFNAMDSVFIISGKLGME--SRGM  
 BT3130 -----NGGHPFTEGNAWQYFVYVPHNIQALMELTGKFAFEQKLDFTFTSTY--K--S-----  
 BT3527 -----APTEGNSWHYTSVFDHPQGLIDMLMGKKEFNAMDSVFIISGKLGME--SRGM  
 BT3773 -----HGVGDFCEGNWQYTFAPQDPYGLISLFGGDKPFTSKLDSVFTNDSM--G-----  
 BT3784 -----EFCENSWHWSFCVFDHPQGLIDMLMGKKEFNAMDSVFIISGKLGME--SRGM  
 BT3858 GA---RGYEGNNGWYRNDVPHNADILSLMGNEQFIANLDRTPTEPLGRSKY-----D  
 BT3962 -----APTEGNSWHYTSVFDHPQGLIDMLMGKKEFNAMDSVFIISGKLGME--SRGM  
 BT3963 GTI-----EGNSWHYTSVFDHPQGLIDMLMGKKEFNAMDSVFIISGKLGME--SRGM  
 BT3965 GSVHGWGDIETGFTWYTWYVQDVQVYVINEA--KELFRKRLDELFTVPLDPI-----  
 BT3990 -----APTEGNSWHYTSVFDHPQGLIDMLMGKKEFNAMDSVFIISGKLGME--SRGM  
 BT3991 -----HRSDDYCEGTAWQYTFVPHDYGVLAKLGGEDAFVEKLDLPTVDSLDG-----  
 BT3994 -----HICESNGWQYFWSVQHDIDGLIDLTGKGNRFAEKLSMPTIYHPAA--D-----  
 BT4073 -----LTHEWYGCIESNAYQGFVPHDVGMLVGGKEEVIANLNLFDHTPSDML-----  
 BT4092 -----PFCENAWHYTSVFDHPQGLIDMLMGKKEFNAMDSVFIISGKLGME--SRGM  
 BT4093 -----GEDFRETYKNYFVYVPHDYGVLAKLGGEDAFVEKLDLPTVDSLDG-----  
 \* : . : \* : . :

```

BT1032 -----YDPANEPDIAYAHLSFYFKGEEWRTQKETORLLDKYFTTKPDGIPG
BT1769 FYAKLPDHTGNVGFQSMANEPVSLHVPVLYNY-AGQPWKTQKRIRQMLKTFWRNLMGIPG
BT1878 -ETTQADVTGLIGQYAHGNEPSSHVTVLYAL-AGRPERTQELIREIFDTQYKPKDGLCG
BT2199 IHEIAEMVALNMGQYAHGNQPVQHAIYLYDY-IGQPWKTQYHLRNVMKLYNSGSKGYCG
BT2629 IHMREMQVMNMGQYAHGNQPIQHMYLYNY-SSEPWKAQYVIREIMNKLYTAGPDGYCG
BT3130 -EQMNHNASGFVGGYAHGNEPSSHVAYLYNF-AGQPWKTQYVSHLNTLYNNTSSGYAG
BT3527 IHEIREMQVMNMGQYAHGNQPIQHMYLYNY-SGEPWKAQYVIREIMNKLYTAGPDGYCG
BT3773 -EGASSDITGLIGQYAHGNEPSSHIAVLYTY-AGEPWKTAEKVRFIMSDFTDQPDGIIIG
BT3784 IHMREMQVMNMGQYAHGNQPIQHMYLYNY-SGEPWKAQYVIREIMNKLYTAGPDGYCG
BT3858 FYHQLPDHTGNVGFQSMANEPVCLHVPVLYNY-AGQPWKTQKRIRVLLNQWFRNLMGVPF
BT3962 IHEIREMEIANMGYAHGNQPIQHMYLYNY-AGEPWKAQYVIREIMNKLYTAGPDGYCG
BT3963 EHNEDIETECLVGGYVHGNPSSHVAYLYNY-TSQPWKTQYVIREIMNKLYTAGPDGYCG
BT3965 GAH---DIQGRIGAYVHGNPSSHVAYLYNY-LKEPWKQKWRITIVDRFYGNTPDALSG
BT3990 IHEIREMTVMNMGYAHGNQPIQHMYLYNY-AGQPWKAQYVIREIMNKLYTAGPDGYCG
BT3991 -ETTSADISGLIGQYAHGNEPSSHVILYLYNY-VNRPWKTQELVDSVYRSQYANNVDGLSG
BT3994 -EELPIFSTGMIGQYAHGNEPSSHVILYLYNY-VGQONLTQKYVAKVMNELYKNEPAGLGG
BT4073 -----WNDYVNHANEPVHFVPLFMQ-LDVPWYTRKWTYICKNAYANKVEGLVGG
BT4092 IHMKEMELANMGQYAHGNQPIQHMYLYCY-AGQPWKTQYVIREIMNKLYTAGPDGYCG
BT4093 -----YNDAMFASGNEPSSHIPWYNYW-IGRPYKTEIIRNVLNEQYSSKIDGLPG
          : . * * *          : :          : :          : . *
BT1032 NDDTGTMSAWAIFNMIIGFYPCDGLPEYALTPVFNKVTIRLDPKMYKENELEVESNRITG
BT1769 DEDEGGMTSFFVFSLLGFYVPTGLPAYTIGSPLETFDAKIRL----SNGAVFEIEAKNAS
BT1878 NDDCGMSAWMYFSAMGFYVVDVPSGDYVFGAPLQPKVILHL----ADGKTFVIAENLS
BT2199 DEDEGGTSAWYVFSAMGFYVPCPGMPEYAGSPLFKKVTLHL----PEGKPFVSAADNA
BT2629 DEDEGGTSAWYVFSALGFYVPCPGTDEYIIGTPLFKSAKLHL----ENKGTITIKADNNO
BT3130 NDDCGMSAWYVFSAMGFYVVPNADGRYIIGSPLLDETLKL----AGNKEFRIRITRKS
BT3527 DEDEGGTSAWYVFSAMGFYVPCPGTDEYILGTPLYFKQMLHL----ENKGTVTISAPNNG
BT3773 NDDCGMSAWYLLSAMGFYQVNPSDGVFAFGSPRFKKIEVKV----RGKVFTEAPNNS
BT3784 DEDEGGTSAWYVFSALGFYVPCPGTQYVILGTPLFKSAKLHL----ENKGTVTIKASNNN
BT3858 DEDEGGMTAFVFSMMGFYVPTPGSPTYNIGSPVFSKAKMEV----GDGHYFEIYAENYA
BT3962 DEDEGGTSAWYVFTALGFYVPCGSNEVYMGAPYFKKATITL----ENKLEISAPKNS
BT3963 NDDCGMSAWYVFSVGMGFYVPCPGTQYVYLGAPYLPYKLTIL----PNGKLEIKAPGVS
BT3965 NDDCGMSAWYVFNCGYVPAVSSNLYNIGSPCAEAITVRM----SNGKNTEMTADNWS
BT3990 DEDEGGTSAWYVFSALGFYVPCPGTDEYVYMGTPPLFKKATLHF----ENGLSVLIDAPNNS
BT3991 NDDCGMSAWYVLLSMGFYVPCGKPVYSIGRPAFDKAVINL----PGERTFVIAKNNS
BT3994 NDDCGMSAWYVFSAMGFYVVPNVSQKYEIGTPLFPEMMLHL----ANGKFTVLAPKVS
BT4073 NEDVQMSAWYVLLAASGHPSCPGNTRMEITSVDFKVEFNLDKSKYHQKVFPTIHAHNN
BT4092 DEDEGGMSWYVLLSLLGIYVCPGTDYVYIGSPVFKKATITM----ENGNKVFIEAPENS
BT4093 NDDLGTMGAWYVFAIGLYPEIPGVGGTVPNTPIFSSVKVHL----KKG---DIVIKGGS
          : * *          : :          : :          : . .          :
BT1032 SETLYINKVLLDGKFKNKYRITHDELVHGKRLIFDLK-----
BT1769 TDNKYIQSATLNGKEWNSWFSHDDLMSGGKLVLMGNKPNKTWASGAEDVPPSLEIK--
BT1878 KEHKYVDSITLNGEYTRNTISHEDILKGGTLVYKMK-----
BT2199 ADRPYIRKALLNGQEFTRNLTLDLQKGGELNLSMDSVFNQQRGTQPADFPYSYSK---
BT2629 LDNRPIKEMKVNKRSQTRNLTLDLQKGGELNLSMDSVFNQQRGTTEKDVPSLSFE--
BT3130 PEDIIYQSATLNGKHKIKOFITHDQIMNOCYVFMKCKPSSGK-----
BT3527 DDKRYISMTLNGKDHTRNLTLDLQKGGELNLSMDSVFNQQRGTTEKDVPSLSFE--
BT3773 KDNIIYQKVVYLNKPYHRSYITVDDIINGSTLKFEMGKPKAKNFKASAMRPIVLNK--
BT3784 TDNRVYKDMKVNKAPTRNLTLDLQKGGELNLSMDSVFNQQRGTTEKDIPIYSLSFE--
BT3858 PDHKYIQSATLNGKTRNLTLDLQKGGELNLSMDSVFNQQRGTTEKDVPSLSFE--
BT3962 DANRYIRSLNNGKNTYKYNLHPDLKGGELNLSMDSVFNQQRGTTEKDVPSLSFE--
BT3963 DKRRYVQSLKLNKESYDKMYITHEIDILKGGVLEFKMSAPNKRKRVSVQDKPYSLTNGIN
BT3965 PKNLVYKELVYNGKDYKSYLYTDDIRDGKLVRFVMSGKPNYKRAVSDAVPPSISLPEK
BT3990 TENFYIDSMSFNGADHTKNYLRRHEDLFGGTTIKVDMSNRPNLNKGTKEEDMPSFSKE--
BT3991 KSNKYISVLLNKGKPLDTPFFGHQDVIAGGVMEIRMTDHPQWGVQR-----
BT3994 KENIYIQSIVKVDGQYKNTYLTHEQIMSGTTEVEMGNTPLEVEFEEOQ-----
BT4073 TNNLYIQKALLNGKYNKYLDFAEIAAGGTLELFMGDKPNTVEWGLSNI-----
BT4092 KENLYIQSATLNGRLLDKNYIHYDDIAEGGVLEFKMSAPNKRKRVSVQDKPYSLSKE--
BT4093 EKDIYIKSMKLSYESTWIDWDQLNSGATIEYRTSGKPKDMKWAG--VPPSFP---
          * .          : *          : :          : *
BT1032 -----
BT1769 -----
BT1878 -----
BT2199 -----
BT2629 -----
BT3130 -----
BT3527 KKK-----
BT3773 -----
BT3784 -----
BT3858 -----
BT3962 -----
BT3963 -----
BT3965 TMRKSSIGF
BT3990 -----
BT3991 -----
BT3994 -----
BT4073 -----
BT4092 -----
BT4093 -----

```

**Figure S3** Expanded *Bt* GH92 sequence alignment used to map sequence conservation in Figure S1.