Supporting Information

Title: Role of a SpoVA Protein in Dipicolinic Acid Uptake into Developing Spores of *Bacillus subtilis*

Author affiliation: Yunfeng Li[1], Andrew Davis[1], George Korza[1], Pengfei Zhang[2], Yong-qing Li[2], Barbara Setlow[1], Peter Setlow[1] and Bing Hao[1]

[1]Department of Molecular, Microbial and Structural Biology, University of Connecticut Health Center, Farmington, CT 06030-3305

[2]Department of Physics, East Carolina University, Greenville, NC 27858-4353

Supporting Information List

Table S1. Structural homologs of the *B. subtilis* SpoVAD; nonredundant top-400 DaLi results.

PDB id	**Z-score**	**Rmsd (Å)**	**Aligned residues #**	**Protein**
3H76	24.8	3.3	261	PQS biosynthetic enzyme PqsD
2GYO	24.7	3.3	254	β-ketoacyl-[acyl-carrier-protein] synthase III
3OYT	24.2	3.1	263	β-ketoacyl-[acyl-carrier-protein] synthase I
1WL4	24.1	2.8	258	acetyl-CoA thiolase
3E60	23.8	3.2	271	β-ketoacyl-[acyl-carrier-protein] synthase II
3HHD	23.5	3.2	271	human fatty acid synthase
3EUO	23.1	3.2	264	fungal type III polyketide synthase
1U0W	22.7	3.3	271	stilbene synthase
3A5Q	22.6	3.2	266	benzalacetone synthase
1EE0	22.6	3.2	267	2-pyrone synthase

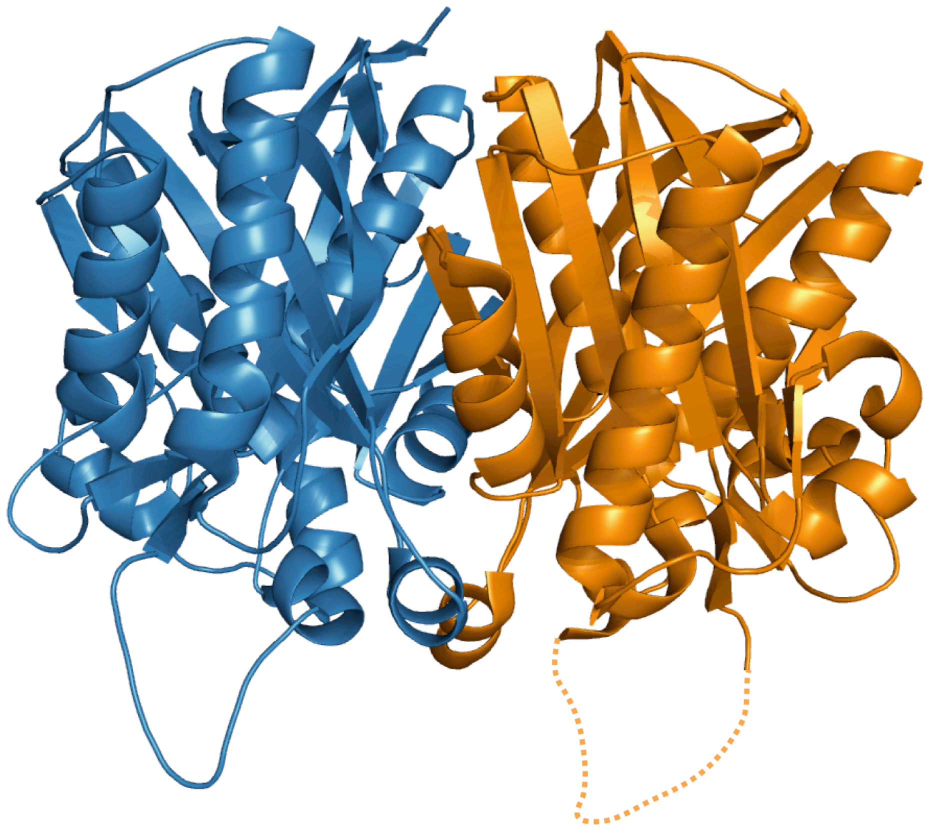


Figure S1. **Ribbon diagram of the SpoVAD crystallographic dimer (PDB code 3LM6).** The two monomers are colored blue and orange, respectively. Dotted lines represent disordered regions.

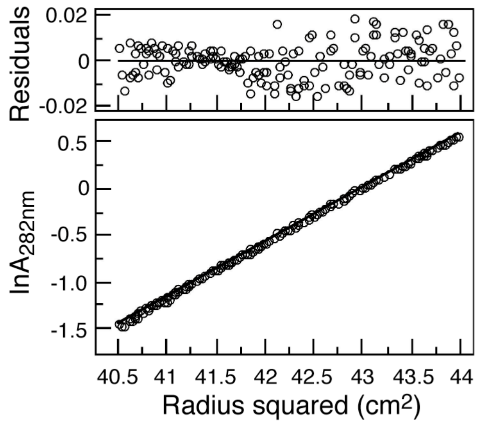


Figure S2. **SpoVAD sediments as a monomer in solution.** Equilibrium sedimentation data (16,000 rpm) of SpoVAD (60 μM) at 20°C in a buffer containing 15 mM Tris-HCl (pH 7.6), 100 mM NaCl and 2 mM dithiothreitol. The data fit closely to a monomer. (Upper) The deviation in the data from the linear fit for a monomeric model is plotted.

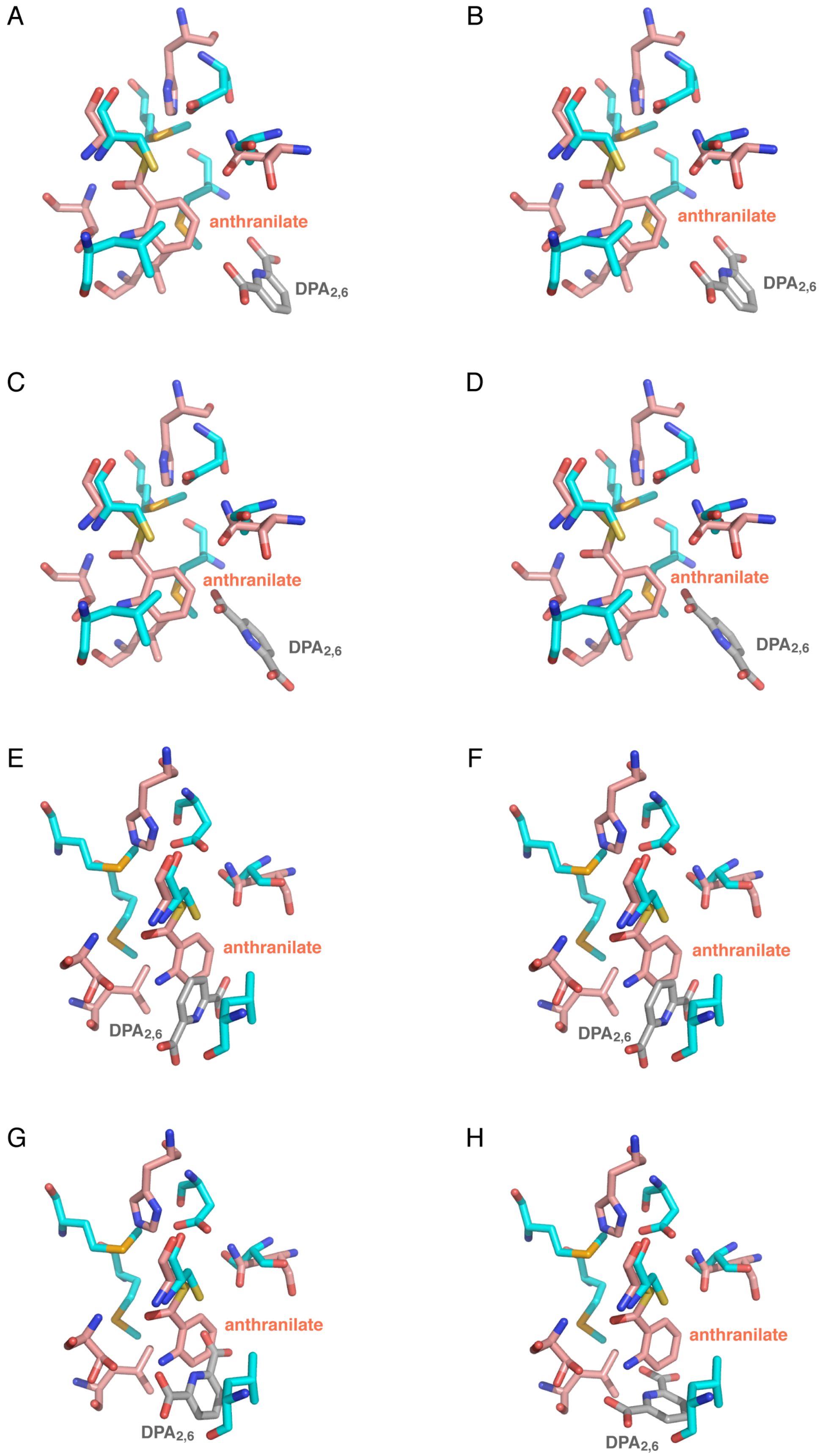
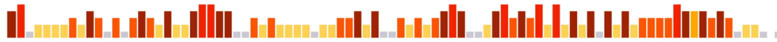

A
anthranilate
DPA2,6
B
anthranilate
DPA2,6
C
anthranilate
DPA2,6
D
anthranilate
DPA2,6
E
anthranilate
DPA2,6
F
anthranilate
DPA2,6
G
anthranilate
DPA2,6
H
anthranilate
DPA2,6

Figure S3. **The docking conformations of the $DPA_{2,6}$ molecule to SpoVAD.** A-H, Eight of the top nine docking solutions for the $DPA_{2,6}$ binding to SpoVAD (PDB code 3LM6) using AutoDock Vina. $DPA_{2,6}$ was docked as a rigid body and all the nine solutions found the $DPA_{2,6}$ molecule in the same presumed binding site in SpoVAD.

```
B. subtilis              1  ---------------------------MKLTGKQTWVFEHPIFVNSAGTAAGPKEKDGPLGSLFDKTYDEMHCNQKSWEMAERQLMEDAVNVALQKNNLT--KDDIDLLLAGDLL
C. saccharolyticum       1  ----------------------------MQIGKASIKFEEPPIIESMASIVGKKEGQGPLGKLFDVVEQDDMFGADTWEKAESALQKQTADLAIEKGDIR--KKDIRYLFAGDLL
S. wolfei                1  -----------------MP-------AKKKRGAQTLVFDQPPVITSWASIAGPKEGQGPWGQDFDWSMEDYLFGEESWEKAENKMLRETVKKALNKRNFD--EADTEILMAGDLL
G. sp. C56-T3            1  ---------------------------MKRVGKQTWVFDDGVYIQAAAAAVGPLEADGPLGSYFDVCHRDLYCGEATWELAERRLMREAVDCCLQKAGVS--AEEVDLFLAGDLL
A. flavithermus          1  --------------------------MMKKIGEQTIVFEKNIYIESSGTVVGPKEREGPLGSSFDRAYDDLYCGEQSWELAERRLMKEAIMICLTKGKVQ--THDVDLFIAGDLL
E. harbinense            1  --------------------------MATRVGRYTIRLDQPPSVMGYASVVGKKEGEGPLADEFDQIEEDTTFGEKTWEKAESRLQNTTLNKALGKADIT--PGTVDIIFAGDLL
C. difficile             1  ------------------MKN-------KRIGKRTVKLENKPTIISTGTIVGPKEGEGPLKDYFDMIMTDDLYGEKTWELAESKMVETASQQAIQKAGKK--LSDVNYMLGGDLI
C. lentocellum           1  ---------------------------MGHIGKQTITFKEPPVITYAYATAGPKEVQGPMGKYFDHPLEDELLGEDTWEKAEARMVQLTIQGLLDRAGKT--PADIHYIYAGDLQ
Paenibacillus JDr-2      1  ----------------------------MLQGKQTWKFENRPVIIGAATVVGPDEGDGPLAGDFDLVHPELDLQQKSWEKAERLLLEQASEFALNNAKIA--KDQLQFYVGGDLM
D. acetoxidans           1  ----------------------------MLRGHQTWVFSSKPCIIASSAIGGPFEAQGALSGDFDILHDDLWLGQTTYEKAEKKLLEDACETAIQKAGLK--KEDIQFFLCGDLM
C. cellulolyticum        1  ------------------MA-------EKHIGKQTVILKNPPSIKATASFVGPKEGKGPLKDDFDYIITDELWGEDSWEKAESKLVRETFAKVLQKAKLS--NDQIDYVLAGDLL
P. polymyxa              1  ---------------------------MRRQGGQTWKFESRPRIIGSATVVGPDEGEGPLSTDFDYIYDNIEINEKTWEKAERRLIEHSTELALINSNLN--KEEVQFFISGDLM
Paenibacillus Y412MC10   1  ----------------------------MKVIGQTWQFENKPVIIGTGTIVGPEEGEGPLANDFDFVYDNIEINEKTWEKAERKLLEDASRKAVEKAGIK--EDQLQFFVGGDLM
T. thermosaccharolyticum 1  ------------------MA-------QKKMGLQTVKFVNPPSIISSGTIVGPKEGQGPLKDYFDMILTDDTYGEKSWEKAECKMFQDSVNLALKKASLN--INDIDYLIGGDLL
P. thermopropionicum     1  ------------------MA------AQKKNGLQTVWFQSPPVIISAASVVGFREGEGPLGKTFDKVVYDSYYGEKTWEKAEQRMMLDAMKKALQKAGLQ--EKDVDFILAGDLL
H. modesticaldum         1  ----------------------------MQQGHQSWRFDNKPVIVASAAVGGPFEARGPLSKDFDLLHGDLWLGQASFEKAERKLLEEACEVAIGKAGLE--KKDIDFFLCGDLM
G. kaustophilus          1  ---------------------------MKRVGKQTWVFDDGVYIQAAAAAVGPLEADGPLGSYFDVCHRDLYCGEATWELAERRLMREAVDCCLQKAGVS--AEEVDLFLAGDLL
E. rectale               1  ------------MRLRVRRVG-----MQQKLGKWSLRFTEAPFILESSSVVGTKEADGPLGGLFDIVGKDDRFGQNTWEKAESSMQRDALTMALGKAHLD--ASDMDYVFAGDLL
C. saccharolyticus       1  ---------------------------MKKVGKSTFKFENCPIISDCFTIVGKKEGEGPLGLFFDMVVEDEYAGQKSWELAEGKLLNIAISKLIQKSGFT--PDEVDLILSGDLL
R. bus                   1  ------------------------------MHKATTIFDQPPSIVTSAAVGGKTENEGPFGKYFDKINDDPYLTADTFEQGENQLQKQAVRIALEKANLK--EDDIDVLIGGDLL
C. kristjanassoni        1  ----------------------------MKLGRSTYRFESNVTISDCFTVVGKKEGQGPLSIYFDMVIEDEYAGQKSWELAEGKLLNIAITKLVQRAGEK--PDNIDLILSGDLL
T. brocki                1  ------------------MA-------EKKLGTQTVKFKNPPSIIAGGTIVGPKEGEGPLRDYFDMILVDDTYGEKSWEKAESKMFQDAVNLVLKKANLK--ISDIDYLLGGDLL
B. brevi                 1  ---------------------------MKRVNRQTWHFDQNVRLQSSAVAVGPKEGEGPLAHLFDKIHDDMYAGQQTWEDAERQMMEDAVNTLLQKASIT--GKDVDLILAGDLL
B. amyloliquefacie       1  ---------------------------MKLTGKQTWVFENPLFINAAGTAAGPKEKDGPLGSLFDKTYDEMHCNQKNWEMAERQLMEDAINSALQKQNLK--KSDIDLLLAGDLL
C. kluyveri              1  ----------------------------MIQGHQSWVFESKPAIIGCSTVGGPFEAKGALAKEFDLLHEDIWIGQDSFEKAEKKLLEEACDLAVKNSKLK--NENINFFVSGDLM
C. owensensi             1  ----------------------------MKLGRATYRFESNVAISDCFTVVGKKEGQGPLSIYFDMVIEDEYAGQKSWELAEGKLLNIAITKLVQREGEK--PDNIDLILSGDLL
C. botulinum             1  ----------------------------MLKGHQSWIFNSKPTILASAAVGGPFEANGALADDFDILYEDLWLGQDSYEKAERALLEHACERVVQKSKIK--KENINFFLSGDLM
B. cellusosilyticus      1  ----------------------------MLQGKRTWLFENKPAIISTGTTGGPFEAKGNIPDDFDILYDDMWIGEDSYEKAERVMFEDACSTAIEKAGLT--KEDIQFMFAGDLV
N. thermophilus          1  ----------------MSSHA-----ARKKIGQQTSYLQYSPTIKGTGTVVGPKEGNGPLAINFHEITQDELLGMSSWEQAESEFLYRACQHALGNAKLD--LNQIDYLFSGDLL
G. sp. WCH70             1  ---------------------------MKLVGKQTWVFEHDIYIQETGTAVGPLEAEGPLRASFDITHDDLYCGAKTWELAERRLMQEAIVVCLQKGNIS--FNDIDIFLAGDLL
T. oceani                1  ------------------MA-------LKKLGKQTVKFENPPSIIATASIVGPVEGKGPLREYFDLILSDVNFQEKSWEKAEKKMLKEAAEMAIRKSGVP--GEQIEYFIAGDLL
A. metalliredigen        1  ----------------------------MLQGRQTWIFDSKPTIVASAAVGGPFEAQGPLREDFDILHEDIWLGQDSFEKAERKLLEQACEIAIDKAGMK--KEDIHFFLSGDLM
M. australiensis         1  ---------------MKFLA-------ANSVGKQTVKLNNAPSVIATASIVGPKEGQGPLGQYFDKVLDDDLYGEKSWEKAEAKFFQEAVEKVIIKGHKD--KQEINYLLGGDLL
D. haf                   1  -------------------M------NLKRIGNQTVVFSNPSVILSSYSVVGPMEGQGPLGKTFNKVWQDNINGNTSWEVAETKMLEEAMEGAVKQSQIK--KDEIDYLLAGDLL
B. thuringiensis.        1  ---------------------------MRLTGKQTWVFQNDIYVNATGTAVGPKEAEGPLGKEFDISYDDLHCGERNWELAERRLMSDSIQQVMQKANVE--TSQIDFFLAGDLL
T. tengconsiensis        1  ------------------MA-------EKKLGAQTVKFKNPPSIIAGGTIVGPKEGEGPLRDYFDMILIDDTYGEKSWEKAESKMFQDAVNLALKKANLK--ISDIDYLLGGDLL
C. thermocellum          1  ------------------MA-------SKRLGKQTVMLQNPPRITSTASIVGPKEGKGPLRLFFDNIIEDEMWGEKSWEKAESKLIRETFSKVLQKGNKT--SQEIDYVISGDLL
L. sphaericus            1  ---------------------------------MVIIFQSKPSLLAGGVVAGPLENRSIFSQYFDAVYDDERWQMKTNEQGHRKMIEEACDAAMKKMKVQ--SGDIDYLMGGDLV
C. phytofermentans       1  --------------MTQTTAK-----QSKMVGKQSVQFSNPPAIIGCASVAGQKEGEGPLGAFFDVIEEDPMFGQKTWEEAESKMLGMAVDVAIHNAGLK--KSDIRYLVGGDLL
B. atrophaeus            1  ---------------------------MKLTGKQTWVFEHPIFVNSEGTAVGPKEKDGPLGSLFDKTYDEMHCNQKNWEMAERQLMEDAVSSALQKNNLQ--KDDIDLFLAGDLL
T. mathranii             1  ------------------MA-------EKKLGSQTVKFKNPPSIIAGATIVGPKEGKGPLKGYFDMILVDDTYGEKSWEKAESKMFQDAVNLALKKANLK--ISEIDYLLGGDLL
B. anthracis             1  ---------------------------MRLTGKQTWVFQNDIYVNATGTAVGPKEAEGPLGKEFDISYDDLHCGERNWELAERRLMSDSIQQVMQKANVE--TSQIDFFLAGDLL
C. novyi                 1  ------------------MT--------KRIGKQTIKLENKPRIIATTSVVGPKEGEGPLKDYFDIILSDDLNGKDSFEKAESSIMYRAVQETLKKSNLE--EKDIHYLVAGDLL
C. hydro                 1  -------------------------------MRQTREFANPPVVLGYATIAGPKEGEGPLATYFDKIINNLYYYEDTWEKAEQKMFLEVVEMALQKASVK--KEEVDFFLAGDLL
C. perfringens           1  ------------------MSN-----DTKRVGSQTIELKSKPRIIATYSVVGPREGQGPLKEYFDKILQDDCNGCESFEKAESSLMITAIEGVIKKGNVK--EEDVDYLFSGDLL
S. lipocalidus           1  -----------------MPR------VSKRTGQQSIYFENPPVITNWANIVGPMEGQGPYGQEFDWVLEDYLFGEQSWEKAEAKMLRETVKLVASKANLP--LGNIEVLLAGDLL
B. cereus                1  ---------------------------MRLTGKQTWVFQNQIFVNATGTAVGPKEAKGPLGKYFDVSYDDLHCGEQNWELAERKLMKNSIEKVVQKGNIM--NSQIDFFLAGDLL
B. megaterium            1  ---------------------------MRKTGKQTWVYENPIYINATGTAVGPKEADGPLGETFDITHKDLHCEENNWELAERRLMTESIDSCLQKVNLS--YEDVDLFLAGDLL
B. weihenstephanensis    1  ---------------------------MRLTGKQTWVFQNDIFVNATGTAVGPKEAEGPLGKDFDISYSDLHCGEENWELAERRLMSDSIQQAMQKGNIK--KSQIDFFLAGDLL
G. sp. Y4.1MC1           1  ---------------------------MKLAGKQTWVFEHDIYIQATGTAVGPLEADGPLRESFDIAHDDLYCGEKTWELAERRLMQEAIGVCLQKGNIS--FSDIDIFLAGDLL
S. thermophilum          1  ------------------MAR-------KHIGAQTVRLEQPPVIVGAAAVAGPVEAQGPLGHEFDLTYPDLTVGQDSFEKAERQMMIDACNLALRRAGIQQKPPAVDYFLAGDLL
A. acidoca               1  ----------------------------MLRGHQSWEFEHQPMILSTAAIGGPMEAKGPLADDFDLLHADIRMEQKSWEQAERTLLDEACQKAIEKAGIT--KDAVSFLIAGDLM
B. tusciae               1  ---------------------------MGRVGHQTYVFERQIGIRTTGVVVGQKEAEGPLAKYFDRLHDDMRMGQKTWEKAERTLLKEAMDLAMRKAELE--KEDIPLYIGGDLM
C. cellulovorans         1  ------------------MG--------IHIGKQTIKLDTRPRIITTTSVVGKKEGEGPLNEYFDIVLQDDLNGTDSYEKAESSIMYKAVTETIKKARLS--EEDIDYLFGGDLL
C. beijerinckii          1  ------------------MKN-------KKIGNQTVKLDNPPKIISTHSIVGPKEGEGPLSKYFDEILNDDMLGKESFEKAESEMMFTAISQALKKGNLK--ETDIDYLFSGDLL
C. bescii                1  ----------------------------MKLGRSTYRFESNVAISDCFTVVGKKEGQGPLSIYFDMVIEDEYAGQKSWELAEGKLLNIAITKLVQRTGEN--PDNIDLILSGDLL
B. pseudofirmus          1  ----------------------------MRVGKTTWQFNQPVYLQSTGTAVGPLEGEGPLGVNFDKVFDSLYCDKENWELAERALMEEAVRISLEKINMK--PEGIDVMLAGDLL
A. oremlandii            1  ------------------MA-------VKKIGSQTIKLLNPPSIISAGSIVGPKEGDGPLAQYFDKIVQDDLLGEDSWEMAESALVKEAANIALGKVNLT--LDDMEYMFGGDLL
C. desulforudis          1  ------------------MP------AAKKTGQQTLVFRNPPVIVVSGTVVGGKEAQGPLGKTFDLVIDDPYHGEKSWEHAERRMLEDAAKTALKKAGLS--QQEIDFFLAGDLE
C. hydrothermalis        1  ----------------------------MKLGRSTYRFESNVAISDCFTVVGKKEGQGPLSIYFDMVIEDEYAGQKSWELAEGKLLNIAITKLVQRVGEN--PDNIDLILSGDLL
B. clausii               1  ---------------------------MGLS-HRTMTFSNPVYIQSSGTAVGPKEGAGPLAGSFDYRYTDIYCDEKSWEMAERALMKQAIKQCLTKNHLS--IDSVDLFVAGDLS
C. kronotskyensis        1  ----------------------------MKLGRSTYRFESNVAISDCFTVVGKKEGQGPLSIYFDMVIEDEYAGQKSWELAEGKLLNIAITKLLQRVGEN--PNNIDLILSGDLL
C.obsidiansis            1  ----------------------------MKLGRSTYRFESNVAISDCFTVVGKKEGQGPLSIYFDMVIEDEYAGQKSWELAEGKLLNIAITKLVQRVGEN--PDNIDLILSGDLL
T. potens                1  -----------MVGTTVVNR------AVKKLGKQTFQLANPPVIVGAAAIVGPKEGQGPLGNIFDLVLPDTLAGEKTWEKAERKMLKDAVNMALQKANLT--TKDIDFFMAGDLL
B. pumilis               1  ---------------------------MKLTGKQSWVFEHKLYVNAEGTAAGPKEKEGPIGHLIDKTYDEMHCDQKNWEMAERKLMEDAISSALSKANLQ--KSDIDLLLAGDLL
A. arabaticum            1  ------------------MN--------KSIGQQTISFESQPKIISSSAIVGPKEGDGPLGEYFDKVLDDSYYQQDTWEKAERKMTKENIDLLLTKAKLS--PDKIDYLVGGDLL
A. degensii              1  -----------------MPAP-------KKIGSQTVAFAQPPVILATAALAGAKEGQGPLRPTFDKVIEDPYYGEDSWEKAERRLLQEALEKALAKANLQ--PEQVDYLLAGDLL
H. orenii                1  ------------------MA--------QKFNGQTLTFQNPPHIISHGSVVGPEEKKGPLGKNFDLTYEDAKCGEKTWEKGEKKMVSDSVNLALQKANLT--TSDIDIILGGDLL
D. reducens              1  ----------------------------MLKGHQTWVFPSKPTILATGTVVGPFEGKGPLANDFDLIHGDLWLGQDSYEKAEKKMLEEACEVAVNKSGLQ--KQDIQFFLGGDLL
B. halodurans            1  ---------------------------MMSPRKQTLKFDKPVYIESTGTAVGPMEGAGPLRHTFDQAYDNLYCGEDSWELAERRLMQDAIEHCLKKVHVE--TERVSFMIAGDLL
T. marianensis           1  MASLTAVASMLEKGRSPRPAAPERRDPGRRVGQQTLVFRTRPRIVAAAAVAGPKEGQGPLGAYWDRTYRDRLLGERSWERAERRMCLDAAVLALEKAGLE--PADIDLLLAGDLM
O. iheyensis             1  ----------------------------MLTGHRTWEFINRPSIITTGTVGGPFEANGSIPEAFDTLHEDMWIKQDSFEKAQQKMLEEACQIAIKKSPIE--KENVEFFIAGDLI
B. licheniformis         1  ---------------------------MKLTGKQTWEFENPLFVNSSGTAVGPKEKEGPLGHLFDKSYDEMHCNQKNWEMAERKLMEDAVQSALSKQNLK--KEDIDIFLAGDLL
Tepidanaerobacter sp. Rel 1 ------------------MS-------IKKLGSQTIRFENPPSIIGAASIVGPKEGQGRFKDYYDLILSDNIYAESSWEKAEKKILKETIEIAIKNANLT--TTKIEYLIAGDLL
M. thermoacetica         1  ------------------MS------APKRVGQHTLQFANPPVIVATASVVGPKEGEGPLGDTFDMVIDDSYFGEETWEKAERKMLEEAVKMVIAKAQLK--PQDIDFLLAGDLL
C. acetobutylicum        1  ------------------MK--------KRLGSQTARIDSKPRIVSAVSMVGPKEGRGPLKDYFDIILSDDMNGEESFEKAESNMLYTVITEVIKKSNLT--ENDIDYLFAGDLL
S. glycolicus            1  -------------------M------NTKRVGNQTVVFQNPPFIISAYSIVGPKEGQGPLAQSFDQILQDTYAGEKSWEKTEQKMIETAIRQTIVKAQFT--NQDIHYMLAGDLL
E. eligens               1  -----------------------------MTGKQSLLFDKDVYVRCAATVVGKKEADGPLGEMFDVAVEDDMFGKKSWEEAESHIQEMAVDKLLIKSGMA--ASDIDYIYAGDLL
G. thermodenitrificans   1  ---------------------------MKRVGKQTWVFDEDIYIRAAATAVGPLEADGPLGGYFDIWHRDLYCGEETWELAERRLMSEAVARCLQKAGVS--AHDVDLFLAGDLL
C. tetani                1  ------------------MS--------KKIGKQTLKLETKPKLISTYSIVGPKEGEGPLNMYFDKIIQDDLAGKDSFEKAESNFLYTAVTESIKNANLQ--EDDIDILFAGDLL
```

Conservation: 100-80-60-40-20%

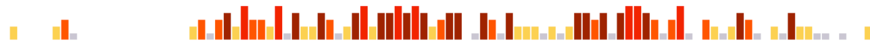
Figure S4 (1 of 4) (Hao)



```
B. subtilis                  87  NQNVTANYVARHLKIPFLCMFGACSTSMETVAVASALVDGGFAKRALAATSSHNATAERQFRYPTEYGGQKPDTATSTVTGSGAVVISQTP------------------------
C. saccharolyticum           86  GQLIATSFGTVDLEIPLFGLFGACSTMGEALNLGAMTVAGGYADKVMAMASSHFATAEKQFRFPLSYGNQRPYSASWTVTGCGAVVLAKHR------------------------
S. wolfei                    90  NQLVSSNYAARELGIPFLGLYGACSTMAESLLLAAILVDGEFFERVIAAACSHHYTAERQFRFPTEQGIQKTPSSQWTATAAGAVLLEPV-------------------------
G. sp. C56-T3                87  NQNATSNYVARELPIPFLCLFGACSTSMQTLAAAAAFVAGGLADRALAATSSHNATAERQFRYPTELGVQKPKTATFTVTGAGAALVGRAP------------------------
A. flavithermus              88  NQTVTANYVASELHIPFLGVFSACSTMTEAMLVASTFIDGAFAHRVLVATSSHYGAAERQFRFPTTIGVQKGESASWTVTGGAAALLSTTR------------------------
E. harbinense                88  NQCIGTSYGLREFGIPFAGLYGACSTMAESLAMAAVFVEGGFANVAAALTSSHFCSAERQYRLPLEYGGQRAPSAQWTVTGSGAVILGREG------------------------
C. difficile                 89  NQIVPASFAARELAIPFLGIYGACSTMAEGLCIGSMLIDGGFADLVLSGTSSHYCTAERQFRFPLELGNQKPMTAQWTVTGAGSVLLAPN-------------------------
C. lentocellum               87  NQVIASTFGVKDFNIPFFGLYGACSTMGESITLASMCIAGGMADLVIAGTSSHNCAAEKQFRFPLEYGGQRTETQTWTTTASGFVLLSRCG------------------------
Paenibacillus JDr-2          86  NQIISNSFAARTLGVPNLGVFGACSTSMESLAIASLMVDSGAGQYVLAGTCSHNCTAEKQFRYPTEYGAQKPPTAQYTVTGAGAVVVAA--------------------------
D. acetoxidans               86  NQIVSSSFAARTLGIPFFGLYGACSSSMEGLALAALLLDSKAAKYVLCATASHNASVEKQFRYPTEYGSQKPPTAQWTVTGAGAAILAP--------------------------
C. cellulolyticum            89  NQCIAANYGLRETEVPFLGLYGACSTMAESMSIGSLMVEGGFAENLVCMTSSHFCSAEKQFRFPLELGSQRTPTSQWTVTGSGAVVLSSK-------------------------
P. polymyxa                  87  NQIVSSSFSASKLAIPYLGVFGACSTSMESLALAALIVDSEAGDYVMAGTTSHNCTVEKQFRYPNEYGAQKPPTAQYTVTGSGAAIVGH--------------------------
Paenibacillus Y412MC10       86  NQIISSSFAARQMGVPYIGVFGACSTSMESLAVASMLVDSGGAQYVMAGTASHNCTVEKQFRYPTEYGSQKPPYAQYTITGAGCSVVSH--------------------------
T. thermosaccharolyticum     89  NQIITSSFAARQFNVPTFGLYGACSTMAEGLSIGSMLIDGGFADYVIVTTSSHFSTAERQYRFPLEQGVQRPFTAQWTVTGSGSSLLSST-------------------------
P. thermopropionicum         90  NQIISANFTASSMSVPFLGLYGACSTMYEGMALGAMLIDGGFADRVLVATSSHYSTAERQYRFPTELGIQRPLTAQWTVTGAGAMLLARE-------------------------
H. modesticaldum             86  NQITASSFAARTMAVPYLGLFGACATSMEGLALASLLVNSRAGRYVLCGTSSHNATAEKQFRYPTEYGGQKPPTAQWTVTGAGVAVVAP--------------------------
G. kaustophilus              87  NQNATSNYVARELPIPFLCLFGACSTSMQTLAAAAAFVAGGLADRALAATSSHNATAERQFRYPTELGVQKPKTATFTVTGAGAALVGRAP------------------------
E. rectale                   97  GQSMATTFGVENFRIPWFGVYGACSTCGESLSLAAAFVAAGYARHAVAVTSSHFGSAEKQFRFPNQYANQRPVSATWTVTGSAAFIVGSGAGGNEI-------------------
C. saccharolyticus           87  NQLTSSHFAARDLNIPFIGLYGACSTFALSVGLACVFVDGGFAKNIIAATSSHFCSAEKQFRYPLELGCQRPPTAQWTVTGAASFLIKPED------------------------
R. bus                       84  DQCVGTTYGVRDYDLPFLGIYAACSTMSEGLLLASMLTAGGFADMAMAVTSSHFCTAERQYRMPLNYGGQRPPTAQWTATAGGALAVSSKG------------------------
C. kristjanasssoni           86  NQLSSSHFAARDLDIPFIGIYGACSTFALSVGLASIFVDGGFAKNVIAATSSHFCSAEKQFRYPLELGTQRPPTSQWTVTGAAAFLIKPED------------------------
T. brocki                    89  NQIISANFAARQLNVPHFGLYGACSTMTEGLTLGAMLIDGGYADYIIAATSSHFSTAERQFRYPLEQGIQRPFTSQWTVTGAGATILAST-------------------------
B. brevi                     87  NQNITTNFTAEKLAMPLLGMYGACSTSMLTLANAAALVNAGYANKAIAACSSHNATAERQYRYPTEYGGQKPPSAQWTVTGAGAGLVGIG-------------------------
B. amyloliquefacie           87  NQNVTANYVARHLKIPFLCLFGACSTSMETVAVASALIDGGFANRAIAATSSHNATAERQFRNPTEYGGQKPDTATSTVTGGGAVIISQNP------------------------
C. kluyveri                  86  NQIITSSFAARTLKIPYLGVFGACSSSMEGLALASFIIDNNGAEYVLCGTSSHNAAAEKQFRYPTEYGGQKPPTAQWTVTGAGAAVIGK--------------------------
C. owensensi                 86  NQLSSSHFAARDLDIPFIGIYGACSTFALSIGLASLFVDSGFAKNVIAATSSHFCSAEKQFRYPLELGIQKPPTSQWTVTGAAAFLIKQED------------------------
C. botulinum                 86  NQVISSSFAARTLGIPYFGIFGACSSSMEGLSLAAQLIESKAAKYVITAASSHNAAAEKQFRYPTEYGVQRPPTAQWTVTGAGAAVLGQ--------------------------
B. cellusosilyticus          86  NQLTPTSFAARTLAAPYFGLFGACSTSMEGLALGAYIVNSGGSKYILTGASSHNSATEKQFRYPTEYGAQKPPTAQWTATASGAAVVSD--------------------------
N. thermophilus              93  NQTVSSNLVAKKLQIPYFGLFGACSTMAEGLLLSSIMVDGGMASNVLTATSSHNKGLERQFRFPTEYGGQRPPYAQWTVSGGGAAIISHQ-------------------------
G. sp. WCH70                 87  NQNVTANYVAQTIPTPFLGMFGACSTSMQTVAVGAALIDGGFAETVLAATSSHNATAERQFRYPTELGVQKPETATFTVTGAGALVMGRKR------------------------
T. oceani                    89  NQIIAASFAARELGFPFFGIYGACSTMAEGLGLGAMLVDGGYADNLVVGTSSHFCTAERQYRFPLELGNQRPPTAQRTTTGAGAAVISSK-------------------------
A. metalliredigen            86  NQITSSSFAARTLGMPYLGVFGACSSSMESLALAALIVDSGAGKNALAAASSHNAAVEKQFRYPTEYGAQKPPTAQWTVTGAGAAVVAR--------------------------
M. australiensis             92  NQLISSDFAARQLGIPFFGLYGACSTMSEAMALGAMLVDGGYADNVVCAASSHFSTAERQYRFPLELGVQRPPTAHWTVTGAGAALVSNQ-------------------------
D. haf                       89  NQIISANFAARTMGIPFIGLYGACSTMALSMSVGTMLIDGGFARKVLVGVSSHHDTAERQYRLPTEQGGQRAMSAQWTVTGAGSLVLGQS-------------------------
B. thuringiensis.            87  NQTVTANYVARKWGIPFLGMFSACATSMETLAVGSAFIDGGFANRVLATVSSHNATAERQFRYPTEYGGQKPGTANSTVTGAGSILISKEK------------------------
T. tengconsiensis            89  NQIISANFAARQLNVPHFGLYGACSTMTEGLTLGAMLIDGGYADYVIAATSSHFSTAERQFRYPLEQGIQRPFTSQWTVTGAGATILAST-------------------------
C. thermocellum              89  NQCIAANFGLRESDVPFFGVYGACSTIAESMSIGAMLVDGGFADNVVCITSSHFCSAEKQFRFPLELGTQRPPSAQWTVTGSGGVLISND-------------------------
L. sphaericus                81  NQMTPTNFAARDLEIPFIGLFSACATSISSVIIACLLTELGAANFSLAGASSQHNAIERQFRYPINYGAQKPQTAQWTVTAAGYALIGKHR------------------------
C. phytofermentans           95  GQLIATSFGIENFEIPLIGVYGACSTMGESIVVASMIIDGGFADRTLAITSSHFAGAEKQFRFPLAYGNQRPKAASWTVTGSGAVILGHEPE-----------------------
B. atrophaeus                87  NQNVTANYVARHLKIPFLCMFGACSTSMETVAVASALIDGGFANRAIAATSSHNATAERQFRYPTEYGGQKPDTATSTVTGSGAVVISQTP------------------------
T. mathranii                 89  NQIISANFAARQLNVPHFGLYGACSTMTEGLSLGAMLIDGGYADYIIAATSSHFSTAERQYRYPLEQGIQRPFTSQWTVTGAGATILAST-------------------------
B. anthracis                 87  NQTVTANYVARKWGIPFLGMFSACATSMETLAVGSAFIDGGFANRVLATVSSHNATAERQFRYPTEYGGQKPGTANSTVTGAGSILISKEK------------------------
C. novyi                     88  NQTAASNFAARDVDIPFLGLYGACSTMSESLGVASMLVNAEYADYAVAATSSHFSSAERQFRFPLEMGSQRTPTAQWTVTGSGAMLIARDGIRE---------------------
C. hydro                     83  NQIISANFVAKSLGIPFFGLYNACATFYEGLILGAMLIAGGFAKKVLVGVSSHYATVERQYRNPTEQGVQKTPTSQWTVTGAAAVVLG---------------------------
C. perfringens               91  NQLTSTNFVAREFKIPFFGVYGACSTMAESLSLSSMIMDGDFAKYAIASTSSHFSSAERQFRFPLEYGCQRKEYSQWTVTGSGAMLLGKS-------------------------
S. lipocalidus               91  NQIISANFAARELKIPFLGLYGACSTMAEGIALGAMLIDGGFYSTLTVGVTSHHHTAERQFRFPTEQGAQRLPSSQWTVTGCGCLIIQRD-------------------------
B. cereus                    87  NQMVTANYVARDMGIPFLGMFSACATSMETLAIGSAFIDGGFAKRILATVSSHNATAERQFRYPTEYGGQRPSTATSTVTGAGSIMISTEE------------------------
B. megaterium                87  NQNVTANFVARHNQIPFLCMFGACSTSMETLATGAALVDAGFANRVIASASSHNATAERQFRNPTEYGGQKPDTATATITGAGAALVSKEK------------------------
B. weihenstephanensis        87  NQTVTANYVAREYGIPFLGMFSACATSMETLAIGSAFIDGGFANRILATVSSHNATAERQFRSPTEYGGQKPGTANSTVTGAGSILISNEA------------------------
G. sp. Y4.1MC1               87  NQNVTANYVARTVPAPFLCMFGACATSMQTIAVGAALIDGGFARTVLAATSSHNATAERQFRYPTELGVQKPKTATFTVTGAGAAVVGRKR------------------------
S. thermophilum              91  NQLITTAFSGLELGVPLYGLYGACSTFTEALSLGAKLIDGGFADRVLCAVSSHYGSAERQYRYPTEYGSQRKPYAQHTVTGAGAVVL----------------------------
A. acidoca                   86  NQIISSSFAARTLAMPYLGIFGACSTSMEGLALASLLVNTGSARYVLAATVSHNAAAEKHYRYPTEYGAQKPPTAQWTVTAGAAAIVAPNDG-----------------------
B. tusciae                   87  DQITTTNFAASFLGSMYLGIYSACGTTVEGLALSAMMMESGAVDRVMTGAVSHTYAAEKLYRFPNEYGVQKPPYAQRTVTGAGVALLEPG-------------------------
C. cellulovorans             88  NQLTATSFAVRKLNIPFFGLYGACSTMTESLSLASLMVEGGFANNVIATTSSHFSSAERQFRFPLEYGSQRNPTAQWTVTGSGSMLLQKS-------------------------
C. beijerinckii              89  NQIISSSFAAREFSIPFFGLYGACSTMSESLSLASIIMDGGFAKYVVAATSSHFSSAERQFRFPLEYGSQRPATAQWTVTGSGALILGHE-------------------------
C. bescii                    86  NQLSSSHFAARDLDIPFVGIYGACSTFALSVGLASIFVDSSFAKNVIAATSSHFCSAEKQFRYPLELGTQRPPTSQWTVTGAAAFLIRQED------------------------
B. pseudofirmus              86  NQIVTSNYAARQLSIPFLGMFSACATVMEAVAVGAVLINSNYVSNALVAVSSHHSTAERQFRYPTEFGGQKPETANYTVTGSGAAVLSNQP------------------------
A. oremlandii                89  NQIMSTSFAARDMKVPFFGLYGACSNFTESLSLASMIVDGGYADQVITTTSSHFSSAERQYRFPLESGNQRPLTSQWTVTGAGATIISSV-------------------------
C. desulforudis              90  NQILASSFAACSLGIPHLGLFGACSTFYEGMVVGAVLIDGGFAEKVLVAVSSHCCTAEKQFRFPTEQGVQRSLAQSWTVTGAGAVILASQ-------------------------
C. hydrothermalis            86  NQLSSSHFAARDLDIPFVGLYGACSTFALSIGLASIFVDGGFAKNVIAATSSHFCSAEKQFRYPLELGCQRPPTSQWTVTGAAAFLIRQED------------------------
B. clausii                   86  NQTTASSFVARDLQLPYIGTYSACATSVETLIISGVWVSSGLADLVLAAASSHYGACEKQFRYPTEFAVQKPETAQTTVTGAGAVLLGKKR------------------------
C. kronotskyensis            86  NQLSSSHFAVRDLDIPFVGIYGACSTFALSVGLASIFVDSNFAKNVIAATSSHFCSAEKQFRYPLELGTQRPPTSQWTVTGAAAFLIKQED------------------------
C.obsidiansis                86  NQLSSSHFAVRDLDIPFVGIYGACSTFALSIGLASIFVDGGFAKNVIAATSSHFCSAEKQFRYPLELGTQRPPTSQWTVTGAAAFLIRQED------------------------
T. potens                    97  NQIISANFTARDLGIPFIGLYGACSTMVESLCLGSMLIDGGFANNVAVAASSHYDTAERQYRFPTELGVQRPPSAQWTVTGAGAMVVSGT-------------------------
B. pumilis                   87  NQNVTANYVARELKIPFLCLFGACSTSMESVAIASALIDGGFANRAIAATSSHNATAERQFRNPTEYGGQKPDTATSTVTGSGAVIISQAP------------------------
A. arabaticum                88  NQVVTSNFSAVKFPVPYFGIYGACSTLAEGLILGAVLIDGGFGNRVINFTSSHYQSAERQYRTPNEYGDKYPPYKQWTATGAGAMITSRD-------------------------
A. degensii                  90  NQIVAANFAARALGIPYVGLYGACATIYEGLALGAMLIDGGFADYVLIGSSSHYSTAERQFRFPTEHGSQRPPTAQWTVTGAGAILL----------------------------
H. orenii                    88  NQLITANFVARDYDVPYLGLYGACSTMVQALGLGAALIDGSYADCVLGYCSSHYQTAERQYRNPLEYGVQYPPYKQWTVTGAGSYILGRV-------------------------
D. reducens                  86  NQIISSSFAARTMAVPYLGLFGACSTSMEGLALGSLLIDSKSAKYLICGTGSHNTAVEKQFRYPTEYGGQKPPTAQWTVTGAGVALLGP--------------------------
B. halodurans                87  NQNITANYVARALGIPFLGTFSACATSMETLAVGSIIVHGGFAQRVITATSSHYATAERQFRQPTEFAKQKPDSATYTVTGAGAALVSNRP------------------------
T. marianensis              114  NQLTCANFTAAELDIPFLGLFAACATWAASLGLAAMATDAGYATRVLCTTSSHHDTAERQYRFPTELGNQRPATSQWTATGAAAAVLAAAGTPLPQGWESANGPAGASPQGSESG
O. iheyensis                 86  NQITPTSFAAKTVGTPYLGVFSACATSMESLALAATIINSGGANYILTGASSHNAAAEKQFRYPTEYGGQKPPTAQWTVTGAGCALVAK--------------------------
B. licheniformis             87  NQNVTANYVARHLKIPFLCLFGACSTSMESIAISSALIDGGFAKRALAATSSHNATAERQFRYPTEYGGQKPGTATSTVTGSGAVVLSQQP------------------------
Tepidanaerobacter sp. Rel    89  NQIISAGFAARELEIPFLGIYGACSTLAEGVTLGSIIVDGGFAENVVAAVSSHFCTAERQFRFPLELGNQRTPTSQWTVTGAGAMVLSSK-------------------------
M. thermoacetica             90  NQTISANYAARSLGIPFLGLYGACSTMYEGMALAAMLIDGGFAHHVVAACSSHYDTAERQYRFPTEQGIQRPPTAQWTVTGAGAVLMAPV-------------------------
C. acetobutylicum            88  NQITSSSFAARDAKIPFYGLYGACSTMTESLSIAAMIMDGGFADYTIAATSSHFSAAERQFRYPLEYGSQRKPDAQWTVTGAGAMLLARQ-------------------------
S. glycolicus                89  NQMISANFTARQLALPFIGVYGACSTMALTTALGGMLIDGGFANKVVCGASSHHDTAERQYRFPTEQGIQRPLWGQWTVTGAGAVLLSDA-------------------------
E. eligens                   85  GQLIATSFGLMRYEIPMFGLYGACSTMGEALSLGAMCVNAGYAQNVIAIASSHFASAEKQFRYPLEYGNQRPVASTWTVTGAGAYIVGDKP------------------------
G. thermodenitrificans       87  NQNATSNYVARELPIPFLCLFGACSTSMQTLAAAAAFVAGGLADTALVATSSHNATAERQFRYPTELGVQKPKTATFTVTGAGAALVGRTP------------------------
C. tetani                    88  NQLSSSSFAARDMNIPFIGLYGACSTMSESLAMASMTIDGDFANNVVAATSSHFSAAERQFRYPLEFGSQRAPTAQWTVTGSGAVVLSRE-------------------------
```

Conservation: 100-80-60-40-20%

Figure S4 (2 of 4) (Hao)



```
B. subtilis                178 ------------------GDIQ------------ITSATVGKVSDLGITDPFDMGSAMAPAAADTIKQHFKDLNRTADDYDLILTGDLSGVGSPIVKDILK-EDGYPVGTK--HD
C. saccharolyticum         177 ------------KEG---IAA-------------VTGITTGRMVDMGIKDSMNMGAAMAPAAFHTIQQNFDDFQVDESYYDKIITGDLGQVGRTILLDFMK-NKGHDLETV--HM
S. wolfei                  180 -------------GT---GPR-------------ITAVSIGKVIDMDQKDVSDMGSAMAPAAAHTIKTHFSDLNRPSDYYDLIVTGDLGKVGMAITQELFI-KDGLQPEPN--YS
G. sp. C56-T3              178 ------------------GRWR------------VRAATIGKVVDAGLKNPLDMGAAMAPAAADTIEQHFRDLGVSPGDYDLIVTGDLSRVGSVIVRELLA-ESGYDVTDV--YN
A. flavithermus            179 ------------------SPFR------------IRTATIGRVIDSGVTDPLDMGRAMAPAAFATIRQHLIDCGLSIEAYDVIITGDLGKVGSSELKSLCY-DAGMNLSER--YD
E. harbinense              179 -------------DG----PY-------------ARELTFGRMVDLGIKDQSNMGAAMAPAAADTLQHYFDDTGVSPDRFDLIVTGDLGAVGSNLLRELLQ-KEGVTLGEN--YN
C. difficile               179 -------------GD---GPK-------------VKYVTVGKVIDEGIDDGNNMGPAMAPAAIDTIYSYFEDTNDDPNSFDIIATGDLGTLGKQIAEDFLK-EKGVDISKV--YT
C. lentocellum             178 -------------EG----PK-------------VTSVTPGKIVDLGVTDTFNMGAAMAPAAVDTILAHLEDTNQTPNDFDAIITGDLGNCGYDIAIDIAG-KLGTDLSGV--LQ
Paenibacillus JDr-2        175 ------------SGE---GPV-------------VEYATIGKVIDLGVKDPFNMGAAMAPAAVDTIQAHFNDTGRTPKDYDLIVTGDLASVGHPLAKELLI-RDGVPMDETT-FK
D. acetoxidans             175 ------------KGQ---GPR-------------VKAATIGRVVDMGITDPFNMGAAMAPAAVDTLEAHFRDTGLGFDNYDLIATGDLGRVGNTIAVDLLK-KHGLPVPQDK-FT
C. cellulolyticum          179 -------------GT---GPY-------------IKHVTTGKINDLGITDANNMGAAMAPAAADTIFAHFKDTGLQPDYYDLIVTGDLGKVGKQICQEMLK-QNGLDIESR--FN
P. polymyxa                176 ------------AKS---GTV-------------VDCATIGRVMDMGIKDPFNMGGAMAPAAADTLLSHFRDTGRDPGYYDLIVTGDLASVGLPIAKELLK-KDGFDMNQTE-FN
Paenibacillus Y412MC10     175 ------------KGD---GPA-------------VTYATIGKVVDLGIKDPFNMGAAMAPAAADTITKHFKDTGRQAKDYDLIVTGDLASVGLPIAKDLLQ-QNDVIMDGTV-FA
T. thermosaccharolyticum   179 -------------GS---GPY-------------ITHVTTGKVVDLGMKDANNMGAAMAPAAADTIITHFNDTGFTINDYDLIITGDMARVGKSILMELLN-KDGLNIEDK--YK
P. thermopropionicum       180 -------------GT---GPV-------------ITHATIGRAIDMGQGDPMDMGAAMAPAAADTIMQHFVDTGRRPDYYDLIITGDLGAYGKKLAEQILA-KNGYDIADR--FT
H. modesticaldum           175 ------------EGK---GPR-------------VTGATIGRIVDMGITDPFNMGAAMAPAAVDTITAHFRDFGIFHEEYDLIATGDLGKVGHHIAKDLFA-KHGMALPEEK-FV
G. kaustophilus            178 ------------------GRWR------------VRAATIGKVVDAGLKNPLDMGAAMAPAAADTIEQHFRDLGVSPGDYDLIVTGDLSRVGSVIVRELLA-ESGYDVTDV--YN
E. rectale                 193 ------QAVCDNADSISRKVQ-------------IEGITTGVVVDYGIKDAMNMGAAMAPAACELIVNHLSDFGRKASYYDRIVTGDLGSVGQKILIDLCR-QKGVDISKN--HE
C. saccharolyticus         178 -----------GTQK----PK-------------ITHFTVGKILDYGIKDSNNMGACMAPAAFDTITTHLEDTKRDFTYYDIIVTGDLGYVGRKLLEELFK-KEGKTLSFE--LL
R. bus                     175 -------------SA----PY-------------VRGVTIGRIVDMGVTDANNMGAAMAPAAFSTISEFFTDTDLSPEDFDHIVTGDLGKVGSRLLCEIFE-HEGMDIRSR--HL
C. kristjanasssoni         177 -----------SLQS----PK-------------ITHFTVGRILDFGIKDPNNMGAAMAPAAFDTIMRHFSDTKRSFEYYDMIITGDLGYVGRKLLDELFK-KEGIRFSSE--LF
T. brocki                  179 -------------GK---GPY-------------ITHATTGKVVDLGMKDANNMGAAMAPAAADTIITHFKDTGFTIEDYDLIMTGDMGRVGRDILLELLY-KEGYDIESK--YK
B. brevi                   177 -------------GN---GPR-------------ITYATMGKVVDMGIKDPFDMGTAMAPAAASTIQTHFEDTGRSPQSYDLIVTGDLAAVGYPILKELML-TEGYDMDQV--YN
B. amyloliquefacie         178 ------------------SDIK------------ITSATVGKVSDLGITDPFDMGSAMAPAAADTIKQHLKDLNRTADDYDLILTGDLSGVGSPILKDLLK-EEGYPAGTK--HD
C. kluyveri                175 ------------AND---RPK-------------VTTATIGKVVDMGVSDPFNMGAAMAAAAVDTIEAHFRDLKRKPSDYDIIATGDLGKIGHEIAEDMLK-RDGFEFSQGV-FK
C. owensensi               177 -----------SLQS----PK-------------ITHFTVGRILDFGIKDPNNMGAAMAPAAFDTIMRHFSDTKRSFDYYDMIITGDLGYVGRKLLDELFK-KEGIRFSSE--LF
C. botulinum               175 ------------NGN---GPK-------------ITSATIGRVIDMGISDPSNLGAAMAPAAVDTIEAHFRDLNIDAFHYDLIATGDLGKVGHELANALLE-KHGINMPTHI-FT
B. cellusosilyticus        175 ------------QGM---GPK-------------VTSATIGKVIDMGLSDPYNMGGAMAPAAVDTIEAHFKERNIDPSYYDLIITGDLGKIGQHVAYDLLQ-DHGLHITKEQ-FV
N. thermophilus            183 -------------GN---GPY-------------VSYVTPGKVMDWGVSDAYDLGSAMAPAAYDTLKQHLSDTGQNVTDYDMIVTGDLGIQGTTMFKELCE-NDGINVQDN--HY
G. sp. WCH70               178 ------------------GRWR------------ITAATIGKVIDEGMTNPLDMGSAMAPAAASTIEQHFIDLGVSPEDYDLIVTGDLSRVGSVIVRQLLA-ESGYDVTSN--YN
T. oceani                  179 -------------DQ---NPR-------------ITYMTTGKVIDMGEANPNDMGTAMAPAAVDTIVRHLEDSNRSPEDYDLIITGDLASVGKDIAERLLK-QKGIDISDR--YT
A. metalliredigen          175 ------------DGD---GPE-------------ILWATIGRVVDMGISDPFNMGAAMAPAAVDTIEAHFRDTKLDPSYYDLIITGDLGTVGHHIAGDLLQ-EHGLNISKER-FS
M. australiensis           182 -------------GE---GPY-------------ITMVTVGKVVDMGIKDSNDMGSAMAPAAADTIIAHFIDTGLTPNDYDLIITGDLGRYGKDIADELVK-KAGYDISAK--HF
D. haf                     179 -------------GE---GPR-------------ITAATIGRIVDFGETDANNMGAAMAPAVADTVINHFQDMNRTPEDYDLIISGDLGSVGLNLTQEVLK-KSGLHFQN---FS
B. thuringiensis.          178 ------------------SAIK------------ITAATIGKVQDLGIANPLDMGSAMAPAAAHTIQQHFEDLKRSANDYDLIVTGDLSAIGTPIAKQLLL-EEGYDIGHI--YN
T. tengconsiensis          179 -------------GE---GPY-------------ITHATTGKVVDLGMKDANNMGAAMAPAAADTIITHFKDTGFTIEDYDLIMTGDMGRVGRDILLELLY-KEGFDIADK--YK
C. thermocellum            179 -------------GS---GPY-------------ITHITTGKIVDMGIKDANNMGAAMAPAAVDTIVTHFKDTGFTPESYDLILTGDLGAIGREICVELTR-EQGFNIENI--YN
L. sphaericus              172 --------------E--DYPS-------------IEAATIGKSVDYGMDDPFHMGAAMAPAAFQTIQSHLQQRNQKIYHYDLILTGDLGQLGLKLLKGMLV-ENGVKNEELTLLR
C. phytofermentans         187 ------KDSEFAKQFSPVMIK-------------VKGITIGKIVDYGIKDSLNMGACMAPAACETIAQNLKDFSIQPDYYDRIITGDLGMIGKDILIDLLK-EKGYDISSN--HM
B. atrophaeus              178 ------------------SEIQ------------ITSATVGKVSDLGITDPFDMGSAMAPAAADTIKQHLKDMNRTADDYDLILTGDLSGVGSPIVKDLLK-EEGFSLGTK--HD
T. mathranii               179 -------------GE---GPY-------------ITHATIGKVVDLGMKDANNMGAAMAPAAADTIITHFKDAGFTIEDYDLIMTGDMGRVGRDILLELLY-KEGYDIESK--YK
B. anthracis               178 ------------------SAIK------------ITAATIGKVQDLGIANPLDMGSAMAPAAAHTIQQHFEDLKRSANDYDLIVTGDLSAIGTPIAKQLLL-EEGYDIGHI--YN
C. novyi                   182 -------------GK---FPY-------------ASYFTVGKIKDYGQTDTSDMGSAMAPAAVDTIKQHFEDTGRSPEDYDLIASGDLGNLGKKITEELLK-EYGYNISAN--YI
C. hydro                   171 ------------EGE---GKIK------------IPRVTVGKALDYGMFDPSDMGGAMAPAAFDTLLAHFKETGLSYDYYDLMATGDLARFGWQILVKLLQ-ERNINFGQR--LT
C. perfringens             181 -------------GD---YPY-------------VTHVTTGIVKDYGIKDADNMGAAMAPAAVDTIYKHFVDTGRKPEDYDVIATGDLGAVGKKLTEELLK-EYGYDISKV--YI
S. lipocalidus             181 -------------GV---GPR-------------ITTATVGKVIDLGQSDVNDMGAAMAPAAADTIVAHFRDTGRDPRYYDLIITGDLGKYGLALMESLCQ-KLNCPLGQH--VS
B. cereus                  178 ------------------SDIK------------ITAATIGRVQDLGIANPFDMGSAMAPAAVDTIQQHFEDLGRSAADYDLIVTGDLSGVGAPIAKQMLL-EEGYDLGNI--YN
B. megaterium              178 ------------------GIVR------------LTSATIGRVMDYGIKNANDMGSAMAPAAADTIQRHLEDLNVSPSEYDLILTGDLSGVGSPILKQLLK-DEGYDISTN--HN
B. weihenstephanensis      178 ------------------SAIK------------ITAATIGKVQDLGIANPLDMGSAMAPAAAHTIQQHFEDLRRSATDYDLIVTGDLSAVGTPIAKQLLL-EEGYDLGNV--YN
G. sp. Y4.1MC1             178 ------------------GRWR------------ITAATIGKVIDAGTVNPFDMGSAMAPAAASTIEQHFIDLGASPADYDLIVTGDLSRVGSVIVRQLLA-ESGYNVTDN--YN
S. thermophilum            178 ----------ASAGE---GPR-------------ITYCTTGKVVDWGVADPLNQGAAMAPAAVDTIEAHLRDTGRSVDDYDLIVTGDLAKFGYEMARRLAA-ERNLYLGEN--YR
A. acidoca                 178 ------------KGE---HPV-------------VTRATIGRVVDMGLTDPFNMGAAMAPAALDTLISHFRDFDLPYDYYDLVLTGDLGRVGSRILRDLMI-EHHVDIPQDR-YG
B. tusciae                 177 -----------------AAPR-------------VEAATFGRVVDSGLKDPFDMGSAEAPAAVATIVAHLQDLGRSLDDYGLILTGDLARVGLTFAKQLLE-GQGVTPTER--VE
C. cellulovorans           178 -------------EDTQ-YPY-------------VSYVTTGKVKDYGIVDANNMGAAMAPAAVDTIKQHFIDTGRSPKDYDIIATGDLGIYGRNITNQLLL-ECGYDIGKR--YI
C. beijerinckii            179 -------------GN---YPE-------------ITYVTTGKVKDYGQKDANNMGAAMAPAAADTIVNHFKDTGRKPEDYDVIATGDLGVIGRELADKLIE-DFGYNIRNQ--HI
C. bescii                  177 -----------SLQS----PK-------------VTHFTVGRILDFGIKDPNNMGAAMAPAAFDTIMRHFSDTKRSFEYYDMIITGDLGYVGRKLLEELFK-KEGISFCSE--LF
B. pseudofirmus            177 ------------------SPIR------------VRQATIGQVVDMGVSNPLDMGSAMAPAAAKTLINHFKDTNRDPSYYDLIVTGDLSRVGSSILRKLVD-DEGYTLGNN--YE
A. oremlandii              179 -------------GE---GPY-------------ITYVTTGKVIDFGIKDAANMGAAMAPAAVDTIRAHFTDTGFKPEDYDLIITGDLGIIGHEIAMELLQ-KEGYNMSKV--FK
C. desulforudis            180 -------------GA---GPY-------------ITHATVGKVIDLGQGDPTNFGAAMAPAAADTIVRHFQDTGRPADYYDLVVTGDLGIYGRELAQGLIG-DHGYQVEGR--FT
C. hydrothermalis          177 -----------SLHS----PK-------------ITHFTVGRILDFGIKDPNNMGAAMAPAAFDTIMRHFSDTKRSFEYYDMIITGDLGYVGRKLLEELFK-KEGIKFSTE--LF
B. clausii                 177 ------------------TEIQ------------LTEATIGRVQDLGITDPMSLGAAMAPAAAEVTIGHLHKTNQQVSDFDLIVTGDLAISGSEIYRKLLR-EQGIYITGR--YE
C. kronotskyensis          177 -----------DLHS----PK-------------ITHFTVGRILDFGIKDPNNMGAAMAPAAFDTIMRHFSDTKRSFDYYDLIITGDLGYVGRKLLDELFK-KEGIKFSYE--LF
C.obsidiansis              177 -----------SLHS----PK-------------ITHFTVGRILDFGIKDPNNMGAAMAPAAFDTIMRHFSDTKRSFEYYDMIITGDLGYVGRKLLDELFK-KEGIRFSSE--LF
T. potens                  187 -------------GI---GPA-------------VTHVTIGKVIDLGIKDANDMGSAMAPAAVDTIVAHFKETTRQPSDYDLIISGDLASIGHSLTIQLVK-QAGYDMSKN--FT
B. pumilis                 178 ------------------SPIQ------------ITSATVGRVMDLGITDPFDMGSAMAPAAADTIKQHLEDLGRTVDDYDLILTGDLSGIGSPILKDLLK-EEGIQLGRK--HD
A. arabaticum              178 -------------GS---GPV-------------ITEATPGQVVDLGIKDPQDMGSAMAPAAADTILQHLEDTGRTPNDYDLIVTGDLGEVGSEILAELLK-EKGVDIADS--HD
A. degensii                177 ----------GRQGK---GPR-------------VTHATIGRVIDLGSTDPMNMGAAMAPAAANTIVSHLGDLGRQVTDYDLILTGDLGYYGRELLLALCR-EYGYELGER--HN
H. orenii                  178 -------------GG---STW-------------ITHATFGKVMDLGQKDANDMGGAMAPAAADTLIQHFKDLNRGPQDYDLILTGDLGRTGRKILDSLLV-EKNINLGEK--LQ
D. reducens                175 ------------EGD---GPP-------------VTSATIGKVVDMGLSDPFNMGGAMAPAAVDTIIAHFRDTLDLSPREYDLIATGDLGRVGHTIARDLLV-QHGMEIPEEI-FT
B. halodurans              178 ------------------SKVR------------IESATIGHVVDYGLKNPLDLGSAMAPAAAHTIVTHLEETDQQVDDYDLIVTGDLSRIGSGILRKLLE-EKRVMFKRN--YD
T. marianensis             229 SGGKAGQPAAGRPGS---APRGPGAGGRPRRAVAVTHATFGRVIDMGLTDPFDMGSAMAPAAADTIRQHLQDTGRRPGDYDLIVTGDLARYGHQLLVELLR-EQGYDAEGV--LA
O. iheyensis               175 ------------QGN---GPY-------------VTSATIGKIVDMGMTDPFNMGGAMAAAAADTIEAHLQDRNIDPSYYDLIITGDLGHIGREVSLDYMH-ERNIDINPDQ-YV
B. licheniformis           178 ------------------GGIK------------ITSATVGRVIDLGITDSQDMGSAMAPAAADTIKQHLEDLGRTPDDYDLILTGDLSGVGSPILKDLLK-EEGINVGTK--HN
Tepidanaerobacter sp. Rel  179 -------------QQ---NPR-------------VTFATTGKVIDMGESDSNDMGAAMAPAAIDTIIRHFQDTGKTPDDYDLIITGDLGSIGKEITEELLK-QKGYNMSNK--YA
M. thermoacetica           180 -------------GN---GPR-------------ITHATIGRVLDVGIKDPNDMGSAMAPAAVDTIVRHFQDTGRGPIDYDLIITGDLGRVGHEIATKLLG-EKKYDVSKN--YS
C. acetobutylicum          178 -------------GD---FPY-------------ITFVTAGKVKDYGITSPDSMGEAMAPAAVDTIKQHFKDTGRTPQDYDVIATGDLGRVGKKITEDLLKKEYGYDMSNI--YM
S. glycolicus              179 -------------GN---GPK-------------VTSATLGKVVDLGETDSNNMGAAMAPAVADTLISHFNDLNRQPDYYDLILSGDLGSVGFELLLQLLD-KSGIKIGNN--FG
E. eligens                 176 ----------LDKKK---CVL-------------IKGITTGKIVDYGVKDSMNMGACMAPAAAELIEANFKDLDVDKDYYDAIFTGDLGEIGNRILSELLK-EKSIDIADK--LY
G. thermodenitrificans     178 ------------------GRWR------------IRAATIGKVIDAGLTNPLDMGAAMAPAAADTIEQHFRDLGTSPSDYDLIVTGDLSRVGSVIVRELLA-ESGYDISDV--YN
C. tetani                  178 -------------GN---FPY-------------VTYVTIGKVKDYGITDANNMGASMAPAAVSTIYNHFKDTGRTPKDYDVIATGDLGKVGKRITIELLK-DQGYDLTSN--YI
```

Conservation: 100-80-60-40-20%

Figure S4 (3 of 4) (Hao)

```
B. subtilis                   260 DCGLLIYTPDQ-Q-VFAGGSGCACSAVVTYSHIFKQLREGKLNRVFVVATGALLSPTMIQQKETIPTIAHGVVFE------------RAGGAS
C. saccharolyticum            261 DCGIEMYNSEE-QDTHAGGSGCGCAASTLCSYILPKIQNGTWKRVLFVPTGALLSTVSFNEGETIPGIAHAVVIE------------NMGNLS
S. wolfei                     263 DCGVLLYDDF--QEVYAGGSGCGCSACFLCGPLLKKMSAGEINKLLFVATGALMSPITSFQGESIPAIAHAVAIE------------N
G. sp. C56-T3                 260 DCGLMIYRPDQ-K-VFAGGSGCACSAVVTYGYLLQELARGTFRRLFVVATGALLSQTMVQQKQSIPAIAHGVVIE------------AAGQEPDLPP
A. flavithermus               261 DCGVLIYDTEQ-Q-AFAGGSGCACCAVVTYSHFLEKMRERKIKTMLVVATGALFNGTMVKQNQTIPAIAHAIAVE------------AVT
E. harbinense                 261 DCGLMIYDRDR-QDVHAGGSGCGCAASVLCSHILNNMKSGHLHEVLFMATGVLMSPLSVQQGQSIPSVAHAVHLS------------TSR
C. difficile                  262 DCGIEMFNLKE-QDVHCGGSGCGCSATVFAGYIYDKLRKKEFNKVMLVSTGALLSPTSTLQKQTIPCVAHAVVIV------------NE
C. lentocellum                260 DCGKLMYDDEK-QGTNCGGSGCGCIASIFTGYFYKQLVEKKLKKILIVPTGALMSPGSTQQGNTIPGIAHAVCIE------------MK
Paenibacillus JDr-2           260 DCGLMVFDVEKQK-VQAGGSGCGCSAVVTYGHILKKIANGDLKRVLVVATGALLSPLSFQQGESIPCIAHAVAIS------------GGKE
D. acetoxidans                260 DCGLLIYKED--QPVMAGGSGCACAATVTYGHIFNRMKKGELKKVLIAATGALLSPLTFQQNESIPGICYAVSFE------------YDGEG
C. cellulolyticum             262 DCGVMIFDAEK-QDTHAGGSGCACSAAVLAGHIYNQITQGKLKRVLFVATGALMSPISSQQNESIPGIAHAIAIH------------RTLN
P. polymyxa                   261 DCGLLIYDINKQKKVIAGGSGCGCSAVVTYGHLLKRIEKGELQKVLVVATGALLSPLTTQQGESIPCVAHAVAFE------------AGGKV
Paenibacillus Y412MC10        260 DCGLMIYDMEKQKYVNAGGSGCGCSAVVTYGHILKRMRKGELKKVLVVATGALLSPLSYQQGESIPCIAHAVALE------------MGG
T. thermosaccharolyticum      262 DCGIEIYD-ES-QDVHSGGSGAGCSAVVLNGWLLSQIKNGTFNRVLFIATGALLSPTSTQQGESIPGIAHAVTIS------------RYT
P. thermopropionicum          263 DCGILIYSGS--QDVHAGGSGCGCSAVVTCGYLMEQLKSGRYKRFMGVGTGALMSPCSVQQGETIPGIGHAVVIE------------AR
H. modesticaldum              260 DCGLLIYDLDRQEEVIAGGSGCGCSATVTYGHLLNRMKQGELRKILIVATGALLSPITYQQNDSIPCIAHAVAIE------------RG
G. kaustophilus               260 DCGLMIYRPDQ-K-VFAGGSGCACSAVVTYGYLLQELARGTFRRLFVVATGALLSQTMVQQKQSIPAIAHGVVIE------------AAGQEPDLPP
E. rectale                    286 DCGMKMFDATN-QNTGSGGSGCACSATVLSAYYLPKLRSGELKRILFVPTGALLSPVSFNEGESVPGIAHGVVLE------------AVI
C. saccharolyticus            262 DCGILMFDQKT-QDTGAGASGCAASACIFGGYFYKLMLQKTFKRILIVPTGALLSSSTVQLGESIPVIAHALSIE------------IE
R. bus                        257 DCGLMIYDSEK-QDVHSGGSGCGCAASMLCGYFIPELRSGRIKNILFAATGALLSPTVSQQGESIPSVSHLVWLS------------HEKGSVIK
C. kristjanasssoni            261 DCGILMFDPET-QNTGAGASGCAASACVFGGYLYKLLRERTFERILIVPTGALMSQSTALLGESIPVIAHALAIE------------MV
T. brocki                     262 DCGIEIFD-ES-QDVHSGGSGAACSAVVLNGWLLSQIQNGVYKKVLFMATGALLSPTSSQQGESIPGIAHAITIS------------TTL
B. brevi                      260 DCGLMIYSPD--QDVFSGGSGCASSAVVTYSYIIDQLNRGLLKNVLVCATGALLSPVSTQQGNSIPCIAHAVALE------------GGS
B. amyloliquefacie            260 DCGLLIYTPDQ-Q-VFAGGSGCACSALVTYTHIFNQLKAGSLNRVLVVATGALLSPTMIQQKETIPTIAHGVVFE------------RAGGAS
C. kluyveri                   260 DCGTLIYKKD--QPVFSGASGCACSAIVTYGHFLKEMERKRLGKILIVATGALMSPMSYQQKESIPSVAHAVSIE------------I
C. owensensi                  261 DCGILMFDPET-QNTGAGASGCAASACVFGGYLYKLLRERTFEKILIVPTGALMSQSTALLGESIPVIAHALAIE------------MVQ
C. botulinum                  260 DCGLLIYKKD--QPSFSGGSGCGCSATVTYGHLLNRMRKGELKKILIVATGALMSPISYQQKESIPSIAHAVSIE------------M
B. cellusosilyticus           260 DCGLQIYKSD--QPVQAGASGAGCSAAVAYGHFINKMKQGDFSKMLVVATGALLSPLSYQQKESIPCIAHAVSIE------------M
N. thermophilus               266 DCGVMIFDSN--QDVHAGASGCACSAVVTYSYIYNLFQNNKINRALVMATGALHNPQIVQQNITIPTICHAVALE------------QQKIGDSNG
G. sp. WCH70                  260 DCGLMIYCPDQ-K-VFAGGSGCACCAVVTYGHLLKEMERGAFQKIFVVATGALFSPTMMQQKQSIPTIAHGVVIE------------AAKGES
T. oceani                     262 DCGLLIYSPQ--QDVHAGGSGCACAAVVTCGYLYKEMQKGRYNRILLVATGALLSTTSVQQGETIPCIAHAVALE------------NL
A. metalliredigen             260 DCGLLIYDREK-QPVIAGGSGCASSATVTYGHLMKRMGKGELKRILVVATGALLSPMSYQQKESIPCIAHAVAIE------------M
M. australiensis              265 DCGVEIFS-PS-QDPHAGGSGCGCSACVLAAYLIPHMWDGTYNSMLFLATGALLSPTSSQQGESIPGIAHAVSIH------------RTK
D. haf                        261 DCGVLIYDKS--QDAHAGASGCGCSAVVFAGHIMEKIKSGQYKRVLLVGSGALHSPTSSLQGESIPGIGHAVVIE------------A
B. thuringiensis.             260 DCGLMIYDSGQ-EEVFAGGSGCACSAVVTYGHLLSEMQKGNLQRIFVVATGALLSPMMMQQKETIPTIAHGVVFE------------RVKGE
T. tengconsiensis             262 DCGIEIFY-ES-QDVHSGGSGAACSAVVLNGWLLSQIRNGVYKKVLFMATGALLSPTSTYQGESIPGIAHAITIS------------TILS
C. thermocellum               262 DCGLMIYDREK-QDTHAGGSGCGCAATVFAGFVYKKLMDGTLKNVLFVATGALLSPVSTQQGESIPSIAHAISIT------------TSPER
L. sphaericus                 257 DAGAEFYGQD--ESFQSGASGAGCSAAVFFSYVLQQMRAGSYKRVLLVATGALLSPLSFQQGETIPCTAHAIEIT------------MK
C. phytofermentans            280 DCGIEIFDSEA-QDTHAGGSGCGCSAITLTGYILQKLRKRDWNRVLFVPTGALLSTVSYNEGQSVPGIAHAVMLE------------SC
B. atrophaeus                 260 DCGLLIYTPDQ-Q-VFAGGSGCACSAVVTYTYIFNQLKAGELNRVLVVATGALLSPTMIQQKESIPTIAHGVVFE------------RAGGAS
T. mathranii                  262 DCGIEIFD-ES-QDVHSGGSGAACSAVVLNGWLLSQIRNEIYKKVLFMATGALLSPTSSYQGESIPGIAHAITIS------------TTL
B. anthracis                  260 DCGLMIYDSGQ-EEVFAGGSGCACSAVVTYGHLLSEMQKGNLQRIFVVATGALLSPMMMQQKETIPTIAHGVVFE------------RVKGE
C. novyi                      265 DCGDEIFDKER-QKVNCGGSGCGCSAVVSCGYIYKNILKGKYKKVLLVSTGALMSSTTSLQGESIPGIAHAVSIE------------FGGM
C. hydro                      256 DCGIMIYRPD--QQVNAGGSGTACSALVTAGYVFNELSQGKIHKFLGIGTGALLSQVTSQQGMSIPVIAHAVSFECY
C. perfringens                264 DCGEKIFDNEK-QGTFAGGSGCGCSAVVACGYLYRMLKEKKLKRVLLVSTGALLSSTSALQGESIPGIAHAVAIE------------YI
S. lipocalidus                264 DCGVLIYDID--QDVHAGGSGCGCSAVMLCGPLLNRMQAKEFNRILLVGTGALMSTTSSLQGETIPCIAHAVAIE------------N
B. cereus                     260 DCGLMIYDAKQ-EEVFAGGSGCACSAVVTYGYLLEEMKKGNLQRIFVIATGALLNPMMIQQKETIPAIAHGVVFE------------RVKGG
B. megaterium                 260 DCGLMVYRPDQ-N-VFAGGSGCGCSAIVTYGHILNELKSGNLKKVFMVATGALLSPTMIQQKESIPTIAHGVVLE------------HIGEGDDQWSM
B. weihenstephanensis         260 DCGLMIYNSNQ-EEVFAGGSGCACSAVVTYGHLLREMQKGNLQRIFVVATGALLSPMMIQQKETIPTIAHGVVFE------------SVKGE
G. sp. Y4.1MC1                260 DCGLMIYRPDQ-N-VFAGGSGCACCAVVTYGHLLKEMERGALQKIFVVATGALLSPTMMQQKQSIPAVAHGVVME------------AVKEER
S. thermophilum               264 DCGVMVYDTTK-QNVDSGGSGCASSAIVTCGHLLKEMARGRLRRLLVVSTGSLFSPTSFQQGESIPTIAHAVAVEA
A. acidoca                    263 DAGVLIYGDL--PEVLAGGSGCGCSASVAYGHVQRRLRDGELRRVLVVATGALLSPISYQQKETIPCIAHAVAME------------WPSAACQAGGSAP
B. tusciae                    259 DCGVMIYRED--QGAFAGGSGAGCSAVVMYGWAVNQMRQGRYRRVLLCGTGALHSPTSVKQGDSIPGICHAVSLV------------MDGDG
C. cellulovorans              263 DCGAEIFDQNT-QDTLAGGSGCGCSAVVFSGYIFKRLLSKQLKRVLLVSTGALLSPTSTLQGETIPGIAHAVAIE------------GGESK
C. beijerinckii               262 DCGEIIFDNEK-QNTLSGGSGCGCSAVVFTGYLYKRLMKKEIKRVLLVSTGALMSTTSSLQGETIPGIAHAVAIE------------MNSN
C. bescii                     261 DCGILMFDPKT-QNTGAGASGCAASACVFGAYLYKLLRERTFERILIVPTGALMSASTVQLGESIPVIAHALAIE------------MV
B. pseudofirmus               259 DCGVMIYHQEQ-P-VFAGGSGAGCPAVVTFGYLMKEMERGIIKRLLVVATGALLSPLMIQQKETIPCIAHAVSFE------------LEEDS
A. oremlandii                 262 DCGVEVFDNER-QDTHSGGSGCGCSAVVSCGYLYKELQAKKLNKILIVSTGALLSSTSSLQGESIPSIAHAVTIE------------NKLN
C. desulforudis               263 DCGILIYDHDK-QDTHAGGSGCACSAVVTGGYLLQQIQTGKLQRLLGVGTGALLNADSPKQGETIPGIGHAVVIE------------RR
C. hydrothermalis             261 DCGILMFDPKT-QNTGAGASGCAASACVFGGYLYKLLRERTFERILIVPTGALMSASTVQLGESIPVIAHALAIE------------MM
B. clausii                    259 DCGVMMYHSEQ-Q-VLAGGSGAGCSASVVYGHIYHRLLAGELKRVLVVATGALLSSMTVLQKQSIPAIAHAIVLE------------RRS
C. kronotskyensis             261 DCGILMFDPKT-QNTGAGASGCAASACVFGGYLYKLLREKTFEKILIVPTGALMSQSTVQLGESIPVIAHALAIE------------ML
C.obsidiansis                 261 DCGILMFDPKT-QNTGAGASGCAASACVFGGYLYKLLREQTFERILIVPTGALMSQSTVQLGESIPVIAHALAIE------------ML
T. potens                     270 DCGILIFDPS--QDTHAGGSGCGCSAVVLSSYILGQMKAGKYKRVLAIGTGALLSPTATLQGESIPGVAHAVVMD------------AHSLSGG
B. pumilis                    260 DCGLIIYTPDQ-N-VFAGGSGCACSAVVTFSHIFQEMKAGNLQRVLVVATGALLSPMMVQQKETIPTIAHGVVFE------------RAGGES
A. arabaticum                 261 DCGLMLYNSD--QGVGAGGSGCACSAVVTASYLLPYINNGNFNRIFVVATGSLLSSVTSLQGDSIPCIAHGVVIENNSKSNAVGSIESKVKSEGK
A. degensii                   263 DCGLMIFDRQK-QDVHAGGSGCGCPAVVTCGYLFQEMSAGRLRRVLGVGTGALMSPVTTKQGDTIPAIAHAVVLEA
H. orenii                     261 DCGAMLLTDE--KKYGSGGSGCAASAVVVGSVIIPKIINADINRIIIIGTGALLSSLTVKQGESIPSVAHAVVIE------K---IPGGENNNG
D. reducens                   260 DCGLLIYNQD--QPVQAGGSGCGCAAVVTYGHILNRMKKGELKKVLIIATGALLSPLSYQQNESIPCIAHAVSLE------------MNV
B. halodurans                 260 DCGVMIYSMDQ-P-VFAGGSGAACSAVVSYGYIFEQMKKGKLKKVLIVATGALLNPMMIQQKETIPCIAHAVSLR------------AEV
T. marianensis                338 DCGLMLYDAS--QGVYAGGSGTACSGMVTAGYLLRRMEQGDLRRVLLVCTGCLHSTTTYKQGDTLPTVAHAVALEVDDAAAAGRAGDGGNGSPAAGVRPGGRGGDPA
O. iheyensis                  260 DCGLTIYREG--QPVLSGASGPACSAIVTYGHFLKQMAEGKLKRILMVATGSLHSPLSVNQKDPIPCIAHAVSIE------------SGSDNE
B. licheniformis              260 DCGLMIYTPDQ-Q-VFAGGSGCACSAVVTFAHIFKEIEAGRLNRVLVVATGALLSPTIIQQKESIPCIAHGVVFE------------RAERGNA
Tepidanaerobacter sp. Rel     262 DCGVLIYYPE--QDVHAGGSGCACSAVVTCGYIFKEMLKGQYNKVLLIATGALLSTTSSQQGETIPCIAHAISLE------------NL
M. thermoacetica              263 DCGILIYDHER-QDTHAGGSGCACSAVVFAGHLMGKLNEGRYKRILGVGTGALLSTTATQQGESIPGIGHGVVIE------------G
C. acetobutylicum             262 DCGEKIFNCDE-QDTNSGGSGCGCSAVVDCGYIYKNMLSGKIKKALLISTGALMSTTSNLQGESIPGIAHAVSIE------------FGGMKNE
S. glycolicus                 262 DCGKMIYSLE--QDTHAGGSGCGCSAVVLAGDLLKRLNNGSLRKILLVGSGSLHSPVSSFQGESIPGIGHAITIE------------V
E. eligens                    262 DCGMLIYEGET-K--CSGGSGCGCSAVVLGSCIMDRLIRGEYKRVLFVPTGALLSTVSYNEGKSVPGIAHGVILE------------GKES
G. thermodenitrificans        260 DCGLMIYRPDQ-K-VFAGGSGCACSAVVTYGYLFHELDRGTFHRLFVVATGALLSQTMVQQKQSIPAVAHGVVIE------------AAKQEPDSLP
C. tetani                     261 DCGDEIFNHEE-QGTASGGSGCGASAVVFCGYLYKKLMEGKFKRVLLVSTGALLSTTSSLQGESIPGIAHAISIE------------YGTGGE
```


Conservation: 100-80-60-40-20%

Figure S4 (4 of 4) (Hao)

Figure S4. **Sequence alignment of SpoVAD orthologs in the *Bacillus* and *Clostridium* species.** A ClustalW alignment of 80 SpoVAD orthologs was performed using DNASTAR Lasergene suite 8 (DNASTAR Inc., Madison, WI). Sequence conservation is shown as a bar graph, with red bars indicating identity among SpoVAD homologs.

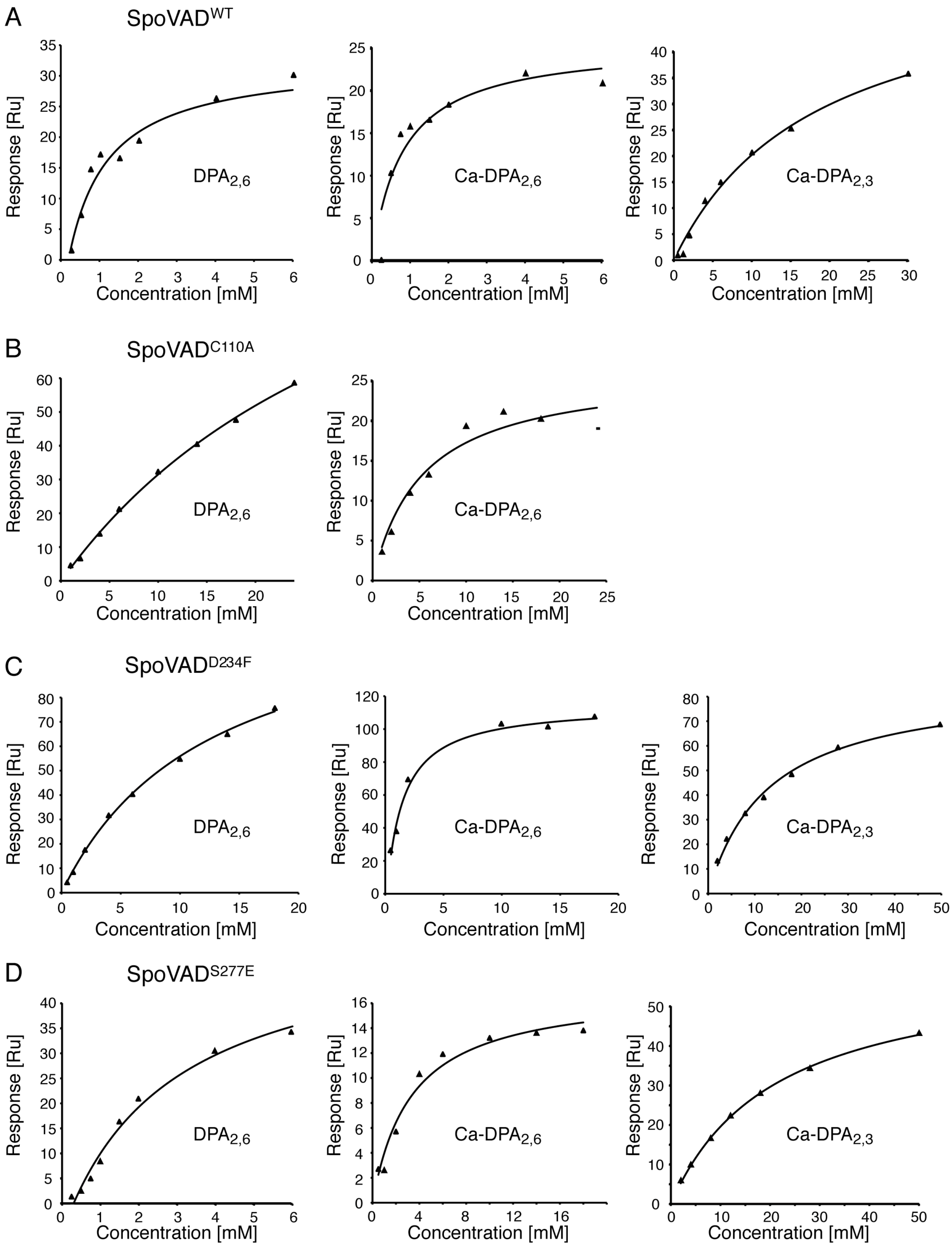

A
SpoVAD^{WT}
Response [Ru]
Concentration [mM]
DPA$_{2,6}$
Ca-DPA$_{2,6}$
Ca-DPA$_{2,3}$
B
SpoVAD^{C110A}
Response [Ru]
Concentration [mM]
DPA$_{2,6}$
Ca-DPA$_{2,6}$
C
SpoVAD^{D234F}
Response [Ru]
Concentration [mM]
DPA$_{2,6}$
Ca-DPA$_{2,6}$
Ca-DPA$_{2,3}$
D
SpoVAD^{S277E}
Response [Ru]
Concentration [mM]
DPA$_{2,6}$
Ca-DPA$_{2,6}$
Ca-DPA$_{2,3}$

Figure S5. **Binding of DPA and Ca-DPA to mutant SpoVAD proteins**. An overlay of the representative Biacore SPR sensorgrams displaying the association curves of 2 mM DPA ($DPA_{2,6}$, $DPA_{2,3}$, $DPA_{2,5}$ and $DPA_{3,5}$; left) and Ca-DPA ($Ca\text{-}DPA_{2,6}$, $Ca\text{-}DPA_{2,3}$, $Ca\text{-}DPA_{2,5}$ and $Ca\text{-}DPA_{3,5}$; right) with the $SpoVAD^{C110A}$ (A), $SpoVAD^{D234F}$ (B) and $SpoVAD^{S277E}$ (C) mutant proteins.

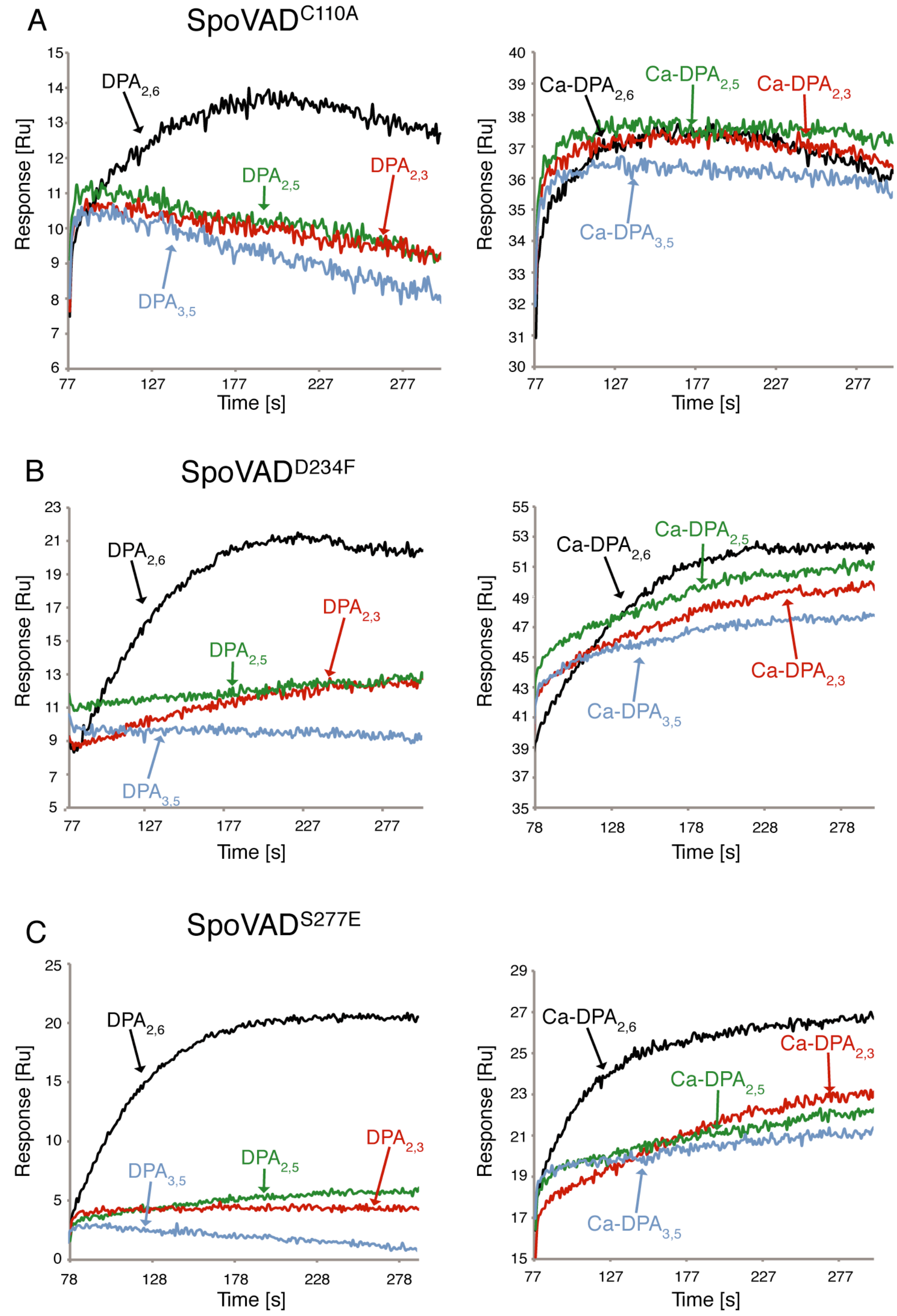


Figure S6. **Measurement of binding affinity of DPA/Ca-DPA to SpoVAD by SPR.** Multiple concentrations of $DPA_{2,6}$, Ca-$DPA_{2,6}$ or Ca-$DPA_{2,3}$ were used to determine their binding affinity (K_D) to the wild-type SpoVAD (*A*), SpoVAD^{C110A} (B), SpoVAD^{D234F} (C) and SpoVAD^{S277E} (D) (also see Table 1). The titration data were analyzed by nonlinear curve fitting to a steady-state model using Biacore T100 analysis software (GE Healthcare, Piscataway, NJ) and OriginPro 7.5 (OriginLab Corporation, Northampton, MA). The binding of Ca-$DPA_{2,3}$ to $SpoVAD^{C110A}$ was too weak to be measured.

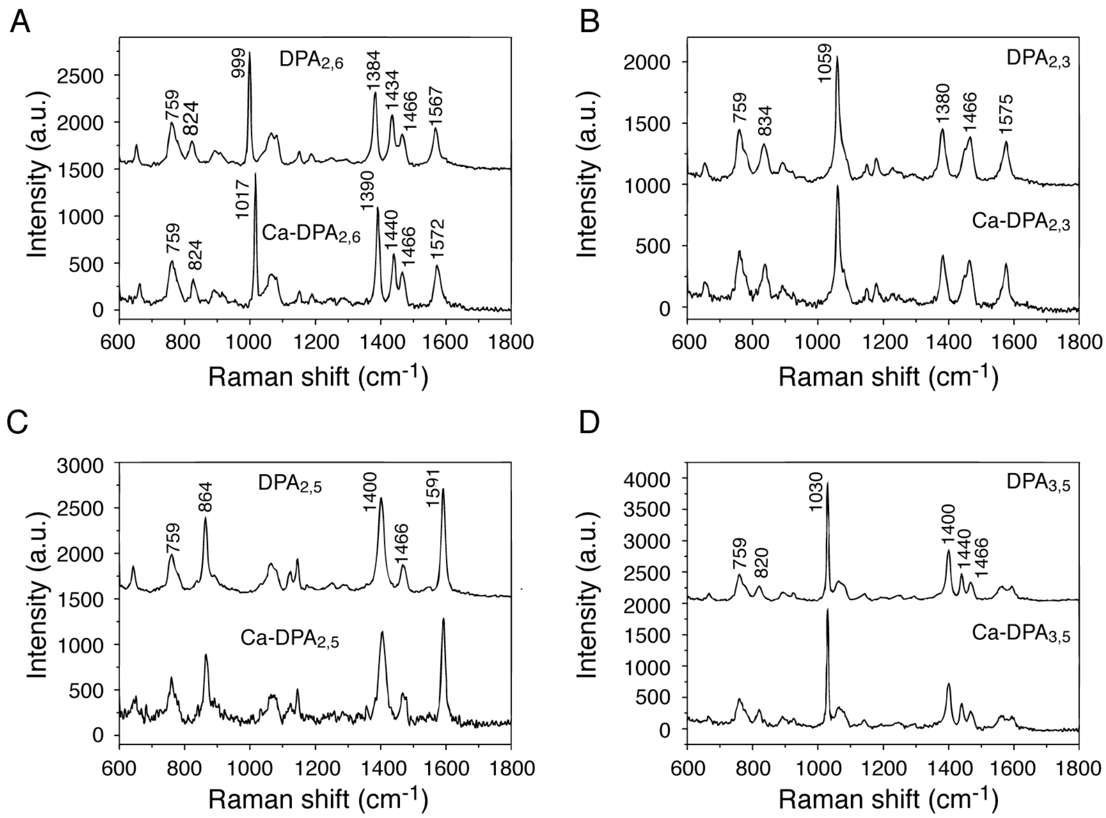


Figure S7. **Raman spectra of $DPA_{2,6}$, Ca-$DPA_{2,6}$ and their isomers.** For each measurement, 60 mM of $DPA_{2,6}$ or Ca-$DPA_{2,6}$ (A), $DPA_{2,3}$ or Ca-$DPA_{2,3}$ (B), $DPA_{2,5}$ or Ca-$DPA_{2,5}$ (C) and $DPA_{3,5}$ or Ca-$DPA_{3,5}$ (D) were used. The laser power used was 30 mW and the integration time for Raman measurement was 20 s.

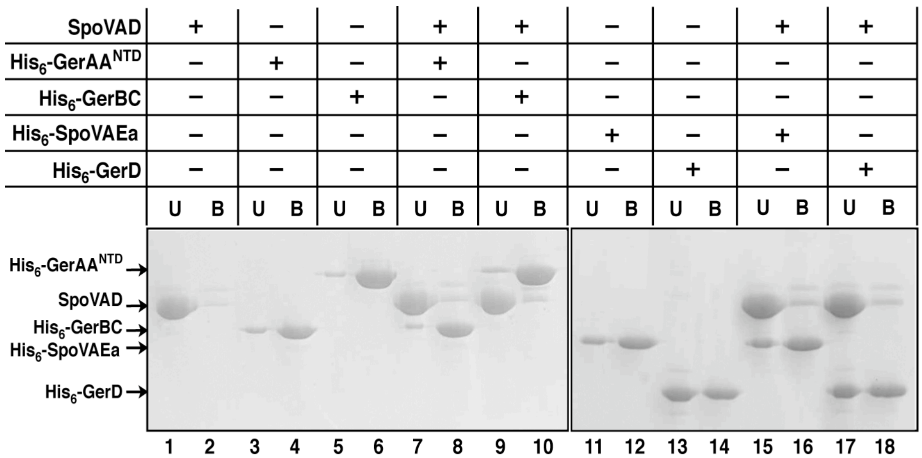


Figure S8. **SpoVAD does not interact with other germinant proteins including SpoVAEa, GerD, GerBC and GerAA^{NTD} *in vitro*.** A Ni^{2+}-NTA pull-down assay was used to characterize the interaction of SpoVAD with the His_6-tagged SpoVAEa, GerD, GerBC and GerAA^{NTD}. Indicated proteins were incubated and reactions were precipitated with Ni^{2+}-NTA resin. The unbound (U) and eluted bound (B) fractions were analyzed by SDS-PAGE and Coomassie staining.