

1 **The engulfosome in *C. difficile*: variations on protein machineries**

2 **Supplementary Information**

3

4 Abigail Kelly¹, Paula S. Salgado^{1#}

5 ¹Institute for Cell and Molecular Biosciences, Faculty of Medical Sciences, Newcastle

6 University, Newcastle upon Tyne, UK

7

8 **# Corresponding author**

9 Address: Institute for Cell and Molecular Biosciences, Faculty of Medical Sciences, Newcastle

10 University, Newcastle upon Tyne, UK. Telephone: +44 (0)191 208 7432;

11 paula.salgado@newcastle.ac.uk

12

13 **Running title: *C. difficile* engulfosome overview**

14 **Keywords:** *C. difficile*, *B. subtilis*, sporulation, engulfment, DMP, Q:AH

15

16 **Conservation analysis – HMM searches**

17 Protein sequences for SpoIIP, SpoIID and SpoIIM from *B. subtilis* (NP_390431.1,
18 NP_391556.1 and NP_390234.1, respectively) and *C. difficile* (YP_001088983.1,
19 YP_001086593.1 and YP_001087714.1, respectively) were used separately to search for
20 orthologues amongst a set of 24 representative endosporeformers [1] using BLASTP [2].
21 Surprisingly, using CdSpoIIM retrieved only two orthologues during this search, despite the
22 sequence conservation revealed in the sequence alignments (Fig. S1). A complete list of all
23 orthologues was generated by combining the sequences retrieved in the BLAST searches
24 and the Bacilli and Clostridia sequences were aligned separately using T-coffee [3]. This
25 alignment was then used to produce a HMM profile in HMMBUILD [4] which was the basis
26 for identifying orthologues of each protein in all Bacilli and all Clostridia as well as all
27 bacteria with JackHMMER [5] within the HMMER3 web server [6]. The alignment of all
28 Bacilli or all Clostridia SpoIID, SpoIIP and SpoIIM homologues retrieved from HMMER
29 searches was then used to generate the final HMM profile from which motif conservation
30 was analysed and weblogs created using Skylign [7]. As the alignment of all bacilli and all
31 clostridial orthologues of SpoIID contained other LytB/PG hydrolysis protein families, as well
32 as true SpoIID orthologues, we focused this particular conservation analysis on the
33 alignment of the representative sporeformers after confirming domain architectures.

34

35 **Signal peptide and transmembrane helix predictions**

36 In order to investigate the presence of signal peptides in *B.subtilis* and *C. difficile* DMP, the
37 proteins sequences were analysed using the SignalP 5.0 [8] and Phobius [9] servers.
38 Similarly, the protein sequences were analysed in the TMHMM [10], Phobius and PSIPRED

39 [11] servers for identification of putative transmembrane helices. Following the standard
40 used for UniProt annotations, a signal peptide was considered to be present if predicted by
41 both SignalP 5.0 and Phobius and a transmembrane helix if predicted by TMHMM and
42 Phobius. If the prediction of the signal peptide and transmembrane covered the same
43 residues, Phobius was used to discriminate between the two possibilities, if necessary.

SpoIIP – representative Bacilli

Bsub_IIP_NP_390431.1 1 MRNKRNRNRIIV.VAVNGG.....KAVKAIPLFIVSLIVIFVLSGVLTSLRPELRLPSSDSFYGIAEELPGDVFADHLLQENHYFASDLS.O.TDSSF.HLSRLSLKLAATSI
 Blicheniformis_IIP_AAU24255.2 1 MRKRGRNRQFV.LAVNGR.....SAVKTVPFLFIVSLLLVLFILSGVLTSLRPELRLPSS.SLYRVADELKGETFGLILGMENHYFASELPE.P.NKRF.ELSPVLVKLATS
 Bhalodurans_IIP_BAB05622.1 1M.FFVTS.YS.....ISVKRLVTFIVATMALMMVLVSE.LVS.NTYFHSQKVRHGMELVTEMLLSMGTENPYFTALP.E.GYEPF.SMATISFEMATSI
 Bclausii_IIP_BAD64188.1 1 MKQGQS.HS.....RFMRIAASLVITLMAVFLCTSVL.VSSGPKTFSSRALANWTTLLSGGHFLKLMGTENRVTEALPDE.ERESF.SLGSLLFAITTSI
 Bcereus_EEK51116.1 1MRKLVLFVITMTLATFFLLISM.VTSMKTKST..YLYNWLNLNLSMNGYMYVLGKENHYFTQEYR.N.LNQDF.SISSFLFSMATNI
 Banthracis_IIP_YP_028184.1 1 MNRGFF.....YVKFT.....SVRKLVLFIATVLTATFFLLISM.VTSMKTKST..YLYNWLNLNLSMNGYMYVLGKENHYFTQEYR.N.LNQDF.SISSFLFSMATNI
 Lsphaericus_IIP_ACA41313.1 1MT.....LLKKL.QWSPGILLFFVFLVPI.....
 Oiheyensis_IIP_BAC13930.1 1M.KKNIKPFIO.....VFSRRSGLYLASVLLVFLMISIL.TTVQPAYRFSSTILSDWTRNIDELVFNMLGMENRAYKSSFP.D.DEELP.KLSSIFFOLTNL
 Gkaustophilus_IIP_BAD76795.1 1 .MKRWRSPNVM.IAIPGA.....SIKKLLMLIILCGMMFMVLGAL.TSLRPEYRPSSSSLNDMAAHPEETFRILFALENRYFAQLLP.K.DRRQT.NYSSLLFRLATSI
 Aflavithermus_IIP_ACJ33207.1 1MMKMM.VTVHGT.....SIKKWLVFIFTFMGLTLVVTATM.TAT.STYRLSSSSVHEVANRPFSTESLVHLLSFENVYFRQTL.P.K.DKQF.PYSQYFFQVTTSV
 Bbrevis_IIP_BAH42982.1 1MATLI.QRQFVV.....LSFITAFLL.FVMTGVL.SLGGNRIASSSAIQQAASHISSLAALGWMCOEIPVLSQVQ.AQADRTNSVTGFLFRLATNI
 Paenibacillus_JDR2_IIP_WP_015843801.1 1 ..MKRTMLFD.LTRGSKRIR.RLLVT.GRTFAVLSSISMFLFVAIGIAGMAQQHSLAAPVSSMSGFAASVSSGFFGDMLEMPQEKSTKS.GEELSGT.QISSFIIRLMTNI

Bsub_IIP_NP_390431.1 102 NLEDPRSFLGRELGFQAFDTEILLAGOG.TDYTNMPESEPPSKVMEEEEREA.....N..LAEIEKO.QTQ..SDNAQKD...PPK.QT.TGD..KKV.V
 Blicheniformis_IIP_AAU24255.2 101 NLKDDPRSFLGRELGFQAFDTEILLAGOG.TDYTNMPESEPPPEVLKKEEREA.....N..LAELEGK.QKK.KTDGK...PPE.QSTGG..RKV.V
 Bhalodurans_IIP_BAB05622.1 91 EMGDPRSLLGRELGFALFDTRIIIVAGEG.TNLSNMPFESPPPEVLLLEEREA.....V..MESSDEE.E.....ADTES...EAE.ROT.D.STR.V
 Bclausii_IIP_BAD64188.1 94 NPDDPRSLLGRELGFSLYDTEIIVAGEG.TDYTSLPVSESSPDDLASEEKA.....K..QALDRK.KAE..EAAKKK...EPD.QANKADEKVVV
 Bcereus_EEK51116.1 82 RFDNDVRSFVKGRELGFQAFDTEIIVAGEG.TNYSNLPVSESSVLEEVKERTS.....G..TGVQPKQ.DTS..KEKK.....QPS.OTTGK..ROV.A
 Banthracis_IIP_YP_028184.1 94 RFDNDVRSFVKGRELGFQAFDTEIIVAGEG.TNYSNLPVSESSVLEEVKERTS.....E..GGQAPK.FDN..KEKK.....OPA.OTTGK..ROV.A
 Lsphaericus_IIP_ACA41313.1 24AGQLP.....NKQASTPIKOPEDKQVFAA.....LEEQNVL.EQT..KEPEPDL...DPF.....K.V
 Oiheyensis_IIP_BAC13930.1 95 QPSDPRSLLGRELGFSLYDQOILMASGD.NHFTSFPVSESNPPELVKEDREA.....V..VEDDVE.E.....LPEE...DIE.OTTGD..RDV.V
 Gkaustophilus_IIP_BAD76795.1 101 NPDDPRSLLGRELGFALYDSKILLAGEG.TDYTNMPESEAPPLEAMLAEREA.....S..VDELEQA.NKN..EDEKPPV...PPS.OTTGG..KKV.V
 Aflavithermus_IIP_ACJ33207.1 95 NLNDPRSLLGRELGFSLYDSEIIVAGEG.TDFTNIPVSEPPLEEVLLAEREA.....A..QEQLQE.E.....TTPAP...APS.OTTGK..KKV.V
 Bbrevis_IIP_BAH42982.1 90 QPGDLRSLLGRELGMVTTEDARFVVQGGKASLADFYLE.YPA.....HPKQVADQ...SQVPIV.EPKPEDTTKSGEAKPAPVN.....SITDG..KKI.V
 Paenibacillus_JDR2_IIP_WP_015843801.1 108 NPGDPRSLLIAGEIFGGMGA.DSSFLIRGCVLGTQPNVWPEENIPIIIGDDTGA.TNGDDHGPNTGGNEPA.....DTP.EPPASPTPTPEKPTSTPPDDAGTHTTGK..DKV.V

Motif 1

Motif 2

Bsub_IIP_NP_390431.1 184 FIYHTNTESYLPPLKKG..ETDPMARHSKANVTLVGDVDFGQALESGGIGATVNTKDIQSKLNKKGKLYARSDSRPVVKDALASNKNLQYIIDIHRDSTRKKDITATKCKS
 Blicheniformis_IIP_AAU24255.2 182 FIYHTNTESYLPPLKKG..EADPDRAIHSKANVTLVSDMLANAMKSGVGAAMVDKTDQANLRKKGWAYARSDSRPVVKEAMAQNKDLQYFIDIHRDSQRKKAITATVCKGS
 Bhalodurans_IIP_BAB05622.1 170 FIHTTBSWESFLPHLPG..VENPNQATHNEMNTIKVGERLGDALKEHGIAQAADKTDMTALLHERNVYRSPAVSRREIVQEAATKEHDDLDFIIDLDRDSVRRDDITVTINGKE
 Bclausii_IIP_BAD64188.1 180 YIGTTBSWESFLPELEMEDIEANANKATHEINMTMADVRLKQALEQHYAQQVEERDVQATLVERGLGYEASDVSREFIKSAKADNDLQYFIDLDRDSIRRRDITVTINGET
 Bcereus_EEK51116.1 161 FIYHTBSWESYLPPLNLTNDPNPNKATSSVSNISILGDRFRFOLAGEGIGATNDKSDVQOKLISKGLNSNSYKMSREIVQOAMAGNKDLQYFIDLDRDSARKNVITKATIGDKS
 Banthracis_IIP_YP_028184.1 173 FIYHSBSWESYLPPLNLTNDPNPNKATSSVSNISIVGDRFRFOLANEGIGATNDKVDVGGOKLISKGLNSNSYKMSREIVQOAMTSNKLEQYFIDLDRDSARKNVITKATIGDKS
 Lsphaericus_IIP_ACA41313.1 74 LVIFTRHESYEPMPVKA..VSGKVAISHETVNTMSLKDKIYHFNVNGKIGTVDVLDTKMLAQGGKGYHOSDVIKRPFLSKHLETNN.YDLIIDLDRDSLKHDITISYNGEN
 Oiheyensis_IIP_BAC13930.1 173 FLYTTRHRESFLPHLPG..VEDPNSAMHAEINISKVSERLQGSLEKRGITVDNTDFGQQLNEQKKYHOSYVSRREAVETAFAFGNKDIQYFIDLDRDALRRDKITIEIEGES
 Gkaustophilus_IIP_BAD76795.1 183 YIYHTBSWESFLPALKG..VTDPLAHPHTVNTIKVGEKLAELKRRGAGVSKDIDVSELLKKGKMYQSQSDMSRQTVVAMKONRDLHYFIDLDRDSRRRKYITATINGVD
 Aflavithermus_IIP_ACJ33207.1 173 YIYHTBSWESYLPALKKG..VTNPNFAHPTVNTIKVGGKLGEELEKRRGAGVVDITDFISKLLKNMYYQSDMSRKTVAAMANNRDVQYLFIDLDRDSRRRKYITATINGVN
 Bbrevis_IIP_BAH42982.1 176 YIYNSHRESWFSSETKSVGT..S.VDHPTRNLSLISKYLSALNDRGIGSDVSTDDIYQOLLNKKMHYSQSAESLQVKAATEKMKLELYFIDLDRDAPRRITVTIKGKT
 Paenibacillus_JDR2_IIP_WP_015843801.1 211 FIYHSBSRESWYPEIADKNG..L.D.ANSASKNITLVGERLAEKLEDLVGVTTSDEPDYPTTVPG..YDWNYSYKYSKKTVEATASNKLLKFFIDIHRDSGRKKITVTKGKS

Motif 3

Bsub_IIP_NP_390431.1 296 YARVAFVVGKKSKNFEENYKIASLELKLMEKKYVGLSLGVFSKGGSPGDNGVYNODLIDRALLLIHEGGVDNMLELQRAANAADVFSEMYWDAEKVNAASGETKK..Q
 Blicheniformis_IIP_AAU24255.2 294 YARVAVFLGKSSNFEENLKLAKELHERMEKKYVGLSRGVISKGAAGDNIYNODLNERSVLLEHGGVDNRELERAAEAMADVFSEMYWNAEKVDADSGEDDKKQ
 Bhalodurans_IIP_BAB05622.1 282 YAQVLFVIGESHDPDYKNELELATKLEKVEKVEYVGLSRGVFKRRDRSSGASTNRYYNODLSDQILIGGGVDNHIDLNRSADAFVFAVDFWILEDG.....
 Bclausii_IIP_BAD64188.1 294 FARPFVVGQNYADFEENLQIANELHKKLEIYVSLSRGVFVKSSGASTNRYYNODLFANSILLIHEGGVDNSLEAYRSIDAFVFAVDFYVNEKEGEN.....E
 Bcereus_EEK51116.1 275 YAKLAFVVGKGNKNYKNEQLQATALHEAINKKYVGVSRGVIQKGFQGTGNGIYNODLGGQAILIHEGGVDNTEELNRSIDALAKAFGEYFWQAEKVN.....
 Banthracis_IIP_YP_028184.1 287 YAKLAFVVGKGNKNYKNEQLQATALHEAITNKKYVGVSRGVIQKGFQGTGNGIYNODLGGQAILIHEGGVDNTEELNRSIDVLAFAFGEYFWQAEKVN.....
 Lsphaericus_IIP_ACA41313.1 185 YAKLAFVVGAEHPNYRNMATAYAESLSTLNAIYVPSKGVISKSGDGDVGRVYNODLAKQMMIILHEGGVNTMEGLYRSADALAEVFAVDFWDAEVS.....P
 Oiheyensis_IIP_BAC13930.1 285 YAKLIFVVGTAHENYENMKTATALNALHEKVEYVGLSRGTEPKKGVGTGNGIYNODLSENAILIHEGGVDNTEELNRSIDVLAFAFGEYFWQAEKVN.....K
 Gkaustophilus_IIP_BAD76795.1 295 YARVAFIIGENATYEKNEQLATALHQLLQKVEYVGLSRGTEPKKGVGTGNGIYNODLSENAILIHEGGVDNTEELNRSIDVLAFAFGEYFWQAEKVN.....K
 Aflavithermus_IIP_ACJ33207.1 285 YARVSLIIGENAKYEKNEQLVTKHEMLKVEYVGLSRGTEPKKGVGTGNGIYNODLSENAILIHEGGVDNTEELNRSIDVLAFAFGEYFWQAEKVN.....K
 Bbrevis_IIP_BAH42982.1 286 YGRVLFVIGKRNKSYKNEAFATLHLEHLMKMYVGLSRGVMEKGAKTDEGEVNSLSPGSLIHEGGVDNTEELNRSIDVLAFAFGEYFWQAEKVN.....K
 Paenibacillus_JDR2_IIP_WP_015843801.1 319 YAQVYFIIGHKNPDWKKNEAFATLHLDVLEKQYVGLSRGILGKTTATGNGEYNOSTLAEDSVLIIHEGGVDNTEELNRSIDVLAFAFGEYFWQAEKVSAS.....K

SpoIID – representative Clostridia

```

Cdifficile630_IID_YP_001086593.1 1 .....MKNPLV.....VLLGFV.TCSVLV.PSLITLVSXKVE.LTEKPESPVINSIKTIKK.SDIEDKGNKEE..KNTVNYETVNKKAPITN...VNHITGKTEKMDMENVLGG
Csubt_tengcongensis_spopt_AAM23469.1 1 .....MRYLLY..G...VLSV.F.LSLIIL.PSIVIFGFSTHEVSV.EK.....N..LVL.....V.E.....GGKSVKLDKLPSEYVIRVFITNQNKIIEEMQLLEDFVFG
Mthermoacetica_spopt_YP_431199.1 1 .....MRKLMGI...FIIILVF.AAVIIT.PVVIIEGIRLFP.P.....VQVQTGKQ..LVRVYFHAGIITKIMPLEDFVFG
Hydrogenoformans_IID_ABB14217.1 1 .....MPNWVS.....IMKKIA.IAILLLVAVTAYTVKQKE.E.....EKVPV.KIQPEKGITLKL...YNHOTDKIETVELEDFVFG
Dhafniense_hp_BAE86697.1 1 .....MVPFPAEQSKH...IVPPLCIVL.SMGLAYERRRD.FVRKE.....L..RILVLLVIIIVLFPWTIV..LRWYNQSEVVRTEDIAIRV..LMPDGGQVKSLEDFVFG
Hthermocellum_IID_ABN53816.1 1 .....MKRFVR..Y...VLMIT.IIVVVL.PMVIVKGCST..VI..E.....D.....IVPEEKEDVK..IKVYVVKDKGVEEEMSLLEDFVFG
Dacetoxidans_IID_ACV64830.1 1 .....MRRKICA.....LILIA.ALLI.....GIPMAVT..L.....MAPV.QIKS.AGTLVRLVYSHEDDQVHEISLEDFVFG
Cacetobutylicum_IID_NP_349464.1 1 .....MVYKKIFM....GLAVTI..IFIFSV.SIFVGGIGDKQNH.TNEN.....SIVKK.SS..YK...ENDDTLNIVKENTKDCD...ISVYMDNEKKIKIEKVNLEDFVFG
Cbotulinum_IID_YP_001252707.1 1 MRRLLNRYTNINKLIVIT.IIS...TV.CIMLIL.SFALSFSENKE.QKVKE.....T..TSKVK.LD..QK.RQEE..KIDINKKYKVKSEPOVRVYFVNEKVKVKSIPLEDFVFG
Cperfringens_IID_BAB81889.1 1 .....MINRKILESVIFIVTII.VSIVFI.PILFGFLSSSS.MAESN.....N..KVE.....KVNKKLVDDVNLKEVEVYISKENKVEKVPLEDFVFG
Aoremlandii_IID_ABW20085.1 1 .....MKGIFP..G...VIFPL.VVTIVM.PMIIVTSCDISIPM..K.....D..K.....PVEKNVVESDITVTVYVNHKTNKIMELELEDFVFG
Sthermophilum_IID_BAD39086.1 1 .....MLRLLIA.....AVLMTA.ALTLAL.PAGIGAGVRMPQP.PDPL.....P..VPVLV.PW..AP...E...PEAPVDEPAGADSLDVKVYFPDRDAIVVMPLEDFVFG
    
```

Motif 1

Motif 2

```

Cdifficile630_IID_YP_001086593.1 97 VLAGMSSEFDIEALKAGAVARTVYVYKQEHGKS...SKHKNAVVCTIDYKHCCQEKYSYDILKLLNGEEW.IKNRYSKIQEAVRGKGOITITYNDAKILP.LYFEST.SS.GKTEEN
Csubt_tengcongensis_spopt_AAM23469.1 79 VVAAMPAEFIEALKAGAVARTYALAKEVALGGKG.CDLHEGVVCTDPEHCCQAWQSVSEELKKGWG.E.NFEKYSKISQAVDSSKGLVMVYQDALIL.PVMHAI.SG.GRTEN
Mthermoacetica_spopt_YP_431199.1 65 VVACMPANFPEALKAGAVARTYTLKKIEAKIKFPDAS.HPNADICTDFHCCQAWAGDVLQRQWGLIG.FWRKKNITQSAVQASGHWMLTYGGQIDIP.VMHAN.GG.GRTES
Hydrogenoformans_IID_ABB14217.1 69 VVACMPANFPEALKAGAVARTYTLKKRILSPQNT..RGYHFGADICTDFHCCQAWYLSGEEELRKRWG.VK.YYVLSRWKAVVSAKGLVLMVNGELIDP.VMHAS.CGGOQKTEA
Dhafniense_hp_BAE86697.1 92 VVAAMPAEFIEALKAGAVARTYAAQRISQSSG...E.AGVVDVDTVNSQVWVISEAKMKKKWLLS.YWRVHSHKIEKAVTSKQDQVLAAGQYID.AFHSSSTGR.KPTEA
Hthermocellum_IID_ABN53816.1 65 VVAAMPADFIEALKAGAVARTYAYGRIKFYSFK.DDTHKQADICTDFHCCQAWTKKEDAMAKWCTLKAFDN.WNKTERAVREKGLIILYDKKVWN.PVHAN.SG.GMTEA
Dacetoxidans_IID_ACV64830.1 90 VVAAMPAEFIEALKAGAVARTYVWRKIDVAGVE..NPAHPCADVCDHHRHCCQAWISEEMKRRWGTVD.FYRYVYKLRKAVRDEKQVWITVYDNVVIDP.AMHAS.CGGOQKTEA
Cacetobutylicum_IID_NP_349464.1 85 VVSSMPVEFIEALKAGAVARTYAVCHMEQFGGRK.YKDAYGADV.VITVQCVYTKQDVLNKKWPKDK.ANEVWNKISTAVEDKQGVITVYDNKIITD.PYVFSVS.GKTEA
Cbotulinum_IID_YP_001252707.1 95 VVSAAMPAEFIEALKAGAVARTYALAHMKGFGENQ.YNKRINADV.DSVOQVFMFKDKRIKSWPKK.RNEVWNKIEEVSVNSKGNVLYVNEVMAPYVYFAT.SN.GKTEA
Cperfringens_IID_BAB81889.1 81 VVSSMPATFIEALKAGAVARTYAVKILITGCNT...IKEGNIC.DITHCCQAVLNIEERKAKWKEG..DEYLLKLLKVVSKPEGKWLNSNDQLVKVPQVST.SN.GMTEA
Aoremlandii_IID_ABW20085.1 70 VVAATPAAPEEALKAGAVARTYVWRKIDVNOAG.HPQHPAIICTNHQCCQEWLSTELERHGNK.NWKQVLPKIEAVSKGLVMTVYMKPIE.PLYHST.SG.GMTEA
Sthermophilum_IID_BAD39086.1 85 VVAAMPADPEEALKAGAVARTYVRRMPQFGGQGGCPLNPEADVCAVYRTSQAYMSRELEAKYGLA.AASEWRRLSQAQAEDEGEVLTQOGELIDA.LYHAV.SG.BMTEA
    
```

```

Cdifficile630_IID_YP_001086593.1 205 SEEVFS.AKYDYLKSVSESPYDKYSPKFASTLKIISNIDFVKSLRRAYSTIVIDV..N.N.LSKQVSTIKRSDAGTVKIKLGNKELTGKDIRTVFKLNSANFDIKFGE.GYIDFVVK
Csubt_tengcongensis_spopt_AAM23469.1 189 SEDVWQ.KKIPYLRVSVSPGEEVASKYKTTVVVSQAEFVKKLLKERQPSLKLTS..NN..ILSEIKNIERTQAGHVTKLIGDVTFFEGKEIKELFGLNSDNTFQSKQ.DDITVITVI
Mthermoacetica_spopt_YP_431199.1 176 AAADVWG.RDVPYLRVSVSPWPKTSPRYSDSRFTSLRYLDSKLGVNLEAVPAAALAAP..GGTAIRVLEKTPQGRVTKIKIGKTFATDRLKLLGLSSTDTWVQGDRIITFHTI
Hydrogenoformans_IID_ABB14217.1 179 AKDVWG.YDVPYLRVSVTCL.ENDY.PVE.TKTFPISYIDKVLGTDLKALAVST..G...TKPVLISERTTGTGRVKKIKLGRIFLAEERYRLLGLASTIMTWTQD.KIIFPTR
Dhafniense_hp_BAE86697.1 199 AEDVWS.SSVPYLRQNVAAAG.EEKPTRYKYTYFTFSDLYQKVGHSSTAKAFTE...SDFVVLSETAAGRAKVVRLVLRGNVYGAQIRTLGLASTIMEITITP.QQIKRITTY
Hthermocellum_IID_ABN53816.1 176 SEDVWEGVEVPYLRVSKSEGEDSPGYKVTYFKEEITKKLKEQYDPMDEP..E..LLEDILIELERTQAGRVKMKIEDVVIKGTIELRALDLRSTNPTIEKKGVEVITTI
Dacetoxidans_IID_ACV64830.1 171 AEDVWK.FKVPYLRVAVPCAGDE.S.KSERTIHIPLEAGKRLGVELQTMVST..E...RSLIAVVKRTAAGYPKVLRIGDKIEPATTVRDKLELSTTKPSWEIAA.GQLSITTT
Cacetobutylicum_IID_NP_349464.1 196 AKEVLG.EDVPYLRVSVSPGEEAHHYKTSKLMSYDGFVDKINSVNNK.INL..IQ..ARVGIISLKRNSGTINQIKVGNVLOATKFRSITGLNSTNPKIKNY.SNIEFDIC
Cbotulinum_IID_YP_001252707.1 206 SENVFN.SEVPYLRVSVKSPGEEKAPKFKSSKLSYDQFKIKIKTQYKCTINK..KD..INKKIIKIKTEFRKILGLNSSKRSLEFKD.NYILVNCN
Cperfringens_IID_BAB81889.1 187 AVAVFS.EDVPYLRVSVQSPGEEISPKYSEISMSISDFKGGKIKKSPNSNLGN..N..INEEVKILSRNKGCTVDDIKIGDVTIKGKEFRKIFGLNSTANFTVLEL.DKINIKCL
Aoremlandii_IID_ABW20085.1 181 SEDVFS.SAIEYLRVSVSPYEEGSPVLVDLTI.SAREFVSKMESRHKDIQINO..RK..LASENLERNTGGSVKRVKVGKELSGRDIRSESFLRSANFTVDVYK.NNVSFSTK
Sthermophilum_IID_BAD39086.1 197 SGDYFA.TDIEYLRVMPVDDTWGAAPRLVDERRFSPEAFVFRALGLGSEEPFLAVQAAARVGGPPVQVARTAGRVKTVAGGMMLTGRVREERLGLRSTDRVYVQD.GEIVVET
    
```

```

Cdifficile630_IID_YP_001086593.1 315 GYGHVGMCSQNCARGMAREGYKYDILSRYYTDTKIKDI.....Y
Csubt_tengcongensis_spopt_AAM23469.1 299 GYGHVGMCSQNCANALAKEGKFKFDILKYYQIEIVKIEDLLKGGH...
Mthermoacetica_spopt_YP_431199.1 288 GYGHVGMCSQNCANGMAREGKNFADILAYVYRGTKIEN.....R
Hydrogenoformans_IID_ABB14217.1 284 GYGHVGLCSQNCANALAKKKNYRILSRYYTPGTVVQY.....K
Dhafniense_hp_BAE86697.1 304 GYGHVGMCSQNCANDLAKAGKTCEBILQHYYPGTQLNLTK...A...
Hthermocellum_IID_ABN53816.1 288 GYGHVGMCSQNCANNLKAGGTYEBILKHYYPGVYLDTTENFEAANP...
Dacetoxidans_IID_ACV64830.1 277 GYGHVGMCSQNCAREFALQGGNYRILSRYYTGVKLVKMT.....N
Cacetobutylicum_IID_NP_349464.1 305 GYGHVGMCSQNCARAMDKGSYKDIIVKHYTGVKIDISY.....
Cbotulinum_IID_YP_001252707.1 316 GYGHVGMCSQNCANSMGKSGKKYDYILKHYKGVNMEILKYS...D...
Cperfringens_IID_BAB81889.1 296 GYGHVGMCSQNCANVMAKESKYYDILKHYFKGSKIEESNEVLI..G...
Aoremlandii_IID_ABW20085.1 291 GYGHVGMCSQNCANGMAKEGKAFNEILKHYQGVSLNILKSHKN..R...
Sthermophilum_IID_BAD39086.1 311 GYGHVGMCSQNCARGMARAGKTYREILSRYYTGVLSRLFEINGENWVPSV
    
```

SpoIID – representative Bacilli

```

Bsub168_IID_NP_391556.1      1 .....MKQFAIT.L.SVLCAL.....ILLVPTLLVIFPFQHN.K.....E.AGAS.....
Blicheniformis_IID_AAU25361.1 1 .....MKQLIIV.L.AGICITL.....ILLIPITLLVLPFFHGG.P.....G.A.....
Bhalodurans_IID_BAB07467.1   1 M.....KRLIVI.G.IVLCV.....ILLLPALLVMPFSPQ.....
Bclausii_IID_BAD66377.1      1 M.....KRIILF.A.AMLIVI.....VLIIVPTVMVALVKDG.EDLTESAF.GLAP.....
Bcereus_IID_EEK47800.1       1 .....MKFSKPLFIT.V.ALLLIAL.....VIIVPAALVIFPAKA.K.....V.GEET.....
Banthracis_IID_YP_031375.1   1 .....MKFSKPLFIT.V.VLLLIAL.....VIIVPAALVIFPAKA.K.....V.GEKA.....
Lsphaericus_IID_ACA38632.1   1 .....MKKWMIS.I.GVLLCLM.....GALYVLPVVLGF.S.....E.....
Oiheyensis_IID_BAC14918.1   1 MQKKYIKRRKPSIQKKIM.MLKQKKQPYASHQEPGLKKIVPFRNRKSRFYQNKSPVPVRLTMLSFF.GILLIF.....ILLVVPATVVNLSKDN.S.....ET.YSQA.....
Gkaustophilus_IID_BAD77625.1 1 MV.....KR.....LKPVMAL.A.LSLFVA.....ILVIPITALVLPFFYDG.K.....V.AKLA.....
Aflavithermus_IID_ACJ35039.1 1 MDERKNDRLQKADGRFLHICLSPRIQFSN.....DKG.....GIGLRQWIIA.T.ITLFTL.....ILLIPALVVLAFGFK.Q.....N.TPLP.....
Bbrevis_IID_BAH46408.1       1 M.....KRYLLMNF.TALPLIL.....LVLMPPALVFPWFSP.....EI.SPVS.....
PaenibacillusJDR-2_IID_WP_015847338.1 1 .....MRYVIGRPKR.....AIR.....RPGSRNLAAV.CWTAEGGIVMSISVLL.....LATSAYRSABSS.....SAGS.DRFQGES

```

```

Bsub168_IID_NP_391556.1      36 .....VESEKT.AVSTKPKASGAETLKASPVSTPVYRTANQSVENIPLLEEVYIGVVASMPAIFKPEALKQAALAARTFIVRLMVSNS.A.VEA.....PKGSLVDDQ
Blicheniformis_IID_AAU25361.1 33 .....IKTEKH.AQQ..HEEKKVTLKESPVSTPVYRTADRKVEEDIPLEEVYIGVVASMPAIFKPEALKQAALAARTFIVRQMVSEQ.A.VQS.....PKGSLVDDQ
Bhalodurans_IID_BAB07467.1   30 .....QMVTEE.TQMAPIEAV.....DLPESELAVHVFRSQTESVAVPLEEVYIGVVASMPAIFKPEALKQAALSLARTFLKQLLEPN.D.IEL.....PEGAMITDE
Bclausii_IID_BAD66377.1      42 .....IEKKQE.AGQTEVRSEQVAAKDTNGDFIAYVRSNAEAVEQIALEDVYMGVVASMPADYEVALKQAALAARTFAVAQMKNGD.KMGDV.....PQGALVTDVY
Bcereus_IID_EEK47800.1       39 .....ASKTPP.....AIESIPAPGKADTAVQVAVYREKQKKVEVLEPVEVYIGVVASEMNASFEHALKQAALAARTFVVQRMLSGG.K.K.....NNADVTDVY
Banthracis_IID_YP_031375.1   39 .....ASKTPP.....AIESIPAPGKVDTAVQVAVYREKQKKVESLPMEEVYIGVVASEMNASFEHALKQAALAARTFVVQRMLSGG.K.K.....NNADVTDVY
Lsphaericus_IID_ACA38632.1   28 .....RREEVQPTAEKACEI..FIEVEGQKKEIPLERVITGVVAEMPVSEFKKALKQAALAARTFALKKTTNYGK.K.A.....IAPVY
Oiheyensis_IID_BAC14918.1   94 .....KENQAA.AQDTEQEVV.....EVGGSPFSVAVMRDKQETVEVDPVEVYAGVVASEMPTFEHALKQAALAARTFTVNLHLHGK.D.N.....NPYDLTDVY
Gkaustophilus_IID_BAD77625.1 40 .....EQLH.....KQEQVQRSAAEGPSIDVAVYRSKEQRVEHIPLEQVYIGVVAEMPVSEFEHALKQAALARTYIVKQLLANQ.P.FRL.....PKGANVTDVY
Aflavithermus_IID_ACJ35039.1 71 .....AEP.....KHVQARDEPIVEVAVYRSKKEKVEHIPLEQVYIGVVAEMPVSEFEHALKQAALARTYIVKQLMHNQ.P.IQL.....PEGANVTDVY
Bbrevis_IID_BAH46408.1       36 .....QPTVA..VPEPAIPAM...KESASSLPVKVYRTEKKAVEVLEPVEVYIGVVAEMPVSEFEHALKQAALAARTYIVRRLKEGKFD.DVP.....SGG.QVLDVY
PaenibacillusJDR-2_IID_WP_015847338.1 63 SGTVSSNKASAD.GFGQAPLPSPGKPNASASSKEYITVYLTKEQRTEERVILEEYVYRGGVVAEMPVDFGLALKQAALAARTYIVRRLLELDD.HSGMPATGAASRADVTDH

```

```

Bsub168_IID_NP_391556.1      132 MFQVYKSKAELKKGWG.TSYETLKLKKTDAVASTCCKILTYNNQPIEASFSTSNCYTENAEAYWTSAPYKSVSPWDK.SP.KYKATKFTAAEFQOKLGVKLDGS...
Blicheniformis_IID_AAU25361.1 127 MFQVYKKNKEELRRLWK.KDYDWKIKKVTAVASTCCKILTYEHKPIDASFSTSNCYTENAGAYWTSAPYKSVSPWDK.SP.KFLNHKFTVAEFOKLGVKLPGS...
Bhalodurans_IID_BAB07467.1   123 MHQVYQSEAEELKERWG.KDYHKKIARIKQAVSETAGQVLTVDGPIITAAFFSTSNCYTENSEDYWKSSEIPYKSVSPWDQ.QHSPRFIGEKVLSVEEFOERLSVTLPPD...
Bclausii_IID_BAD66377.1      139 AHQVYQSKDELKKEKWK.DSFDSNWEKVETAVRDTSCGEIMTYDGEPIITASFSTSNCYTENAEADYWEQEVYKSVSPWDK.Q.SPRFANETVPIPISEFEAALGITLQND...
Bcereus_IID_EEK47800.1       128 KDQVYKKGDELKKGWG.NNYENNLKIEEAVSKTAGQVLTVDGKPIISAFSTSNCYTENAEADYWGNDYPIYKSVSPWDQ.A.SPKFTESEQFTVADFOKRLGVKVLAD...
Banthracis_IID_YP_031375.1   128 KDQVYKSKDELKKGWG.NNYENNLKIEEAVSKTAGQVLTVDGKPIISAFSTSNCYTENAEADYWGNDYPIYKSVSPWDQ.A.SPKFTESEQFTVADFOKRLGVKVLAD...
Lsphaericus_IID_ACA38632.1   105 AKQVYFDESQRKANWA.SNFLGNEKKIVEAINETKQVLLYNNLITAMFHSSTSNCYTENSAAYGYSGNNIPIYKSVSPWDQ.A.SPKFEAEQEWLSLAQWNLKLPVQWQAS...
Oiheyensis_IID_BAC14918.1   184 NHQVYKSPNEELKELWK.DKYEENMKKINSAVKETKQGITHTNDAPITAAFFSTSNCYTENSEADYWEQEVYKSVSPWDK.EE.SPKYTSQSMFTLEEVSAALDELAE...
Gkaustophilus_IID_BAD77625.1 131 AHQVYYSDDDELKRLWG.SDYDWKMKKVTAVMDTTCQGITLYNNEPIEIAFFSTSNCYTENSEAYWQSDFPIYKSVSPWDK.Q.SPKFYQRKTMPPVAEFERRLGVLELPAD...
Aflavithermus_IID_ACJ35039.1 156 MHQVYYSDEQLKELWG.ISYDWKIKKIKKAVEATTCQGITLYNNEPIEIAFFSTSNCYTENSEAYWQNAFPIYKSVSPWDQ.Q.SPKFYEQVTLVSRFEORLGVKLPK...
Bbrevis_IID_BAH46408.1       128 QHVMYMDFKQRRERWG.DQYEWKMKRIQAVMATAQVILTYQNEPIEIAFFSTSNCYTENSEADYWEKEIPIYKSVSPWDQ.I.SPRYEETVVLSTVELEKRLGVKLTQEA...
PaenibacillusJDR-2_IID_WP_015847338.1 174 DHQIYISIDKLELWPKGEGESRMAKLNKAVSETRCGEVITYEGEPIEIAFFSTSNCYTENSEAYWEELPIYKSVSPWDKQLSPRYEETVVKLQLSDFNARLGL.ASKK...

```

```

Bsub168_IID_NP_391556.1      239 SAVGKITGTETPGHQVAVAVINGKTLKGRDIREKLGKLSADFEWK..RNGDITITVTRGFGHGVGMSOYGANFMAKEGKTVDDIVKYYQGTQITSEADAFLNKYMMAK.
Blicheniformis_IID_AAU25361.1 234 NSIGKITERTTPGKRVAVAVINGKTKLEGRDIREKLGKLSADFDWK..RNGDITITVTRGFGHGVGMSOYGANHYMAKDGKKAEDIVKYYKGTAVSNADPFLNKYMMAK.
Bhalodurans_IID_BAB07467.1   231 GSIQSIVARTEGGRVAVKVNINGTISGRVRELELNSADFEWK..RQGNIVIVTRGFGHGVGMSOYGANGMALNGSTYQEIISHYQGVIEISVGPYEQQLMARGD
Bclausii_IID_BAD66377.1      246 GSIQTIARTTAGGRVETVTFENKTLISGREVREALELNSADFSWH..VSNQGVVIVTRGFGHGVGMSOYGANGMAQEGKTYKDIVQHYFKDVEIEISMDELVPT...S
Bcereus_IID_EEK47800.1       235 KGVGNIKDLTEGKRVKDVAFQKTLTCKEVRKLDLSSDFETWK..QGGDKIIVTRGFGHGVGMSOYGANGMAABGKKYTDIVAHYKGVIEIKTMNDYEGKLMVKK.
Banthracis_IID_YP_031375.1   235 GKVDIKGRTEGKRVKDVAFQKTLTGRDVRKLELSSDFETWK..QGGDKIIVTRGFGHGVGMSOYGANGMAABGKKYTDIVAHYKGVIEIKTMNDYEGKLMVKK.
Lsphaericus_IID_ACA38632.1   212 DFNRIQLVYNDISGRVERLQGLGNVWVTRGKRVRELGIPSTDFEISIVYNAKTRVHVTRGFGHGVGMSOYGANGMAABGKTAETAEIISHYQDIEIKKIDACLK.....
Oiheyensis_IID_BAC14918.1   292 QAVPEMIEITRETSGRVKNLVLAGKTIISGRIVRELELSSDFETVE..QKNHLLVFTVEGFGHGIEMSOYGANGMAKEGKTYEDIHYHYKDVIEINSVDETAPTLVS..R
Gkaustophilus_IID_BAD77625.1 238 GSVGVIVSRTPGRVGEVKIGKFTGRDVRERLGLPSDFETWV..RNGDDIITVEGFGHGVGMSOYGANGMAKEGKTYADIVKYYRQVGHKKTATAPLNKLTAKNG
Aflavithermus_IID_ACJ35039.1 263 GSVGVILARTTPGKRVAVLDVINGKTLTGRDVRKLELSSDFETWV..RQGNIVITRKGFGHGVGMSOYGANGMAKQKTYKDIVKYYQGVQIHDISLVASK....F
Bbrevis_IID_BAH46408.1       239 GSWYRIESRTPGNRVGTISIGKEFTGREPREKLNLSSEFTLD..LRGNQVITRKGFGHGVGMSOYGANGMAKSKSAEQIVKHFYQGISIQEYKRVIPA.....
PaenibacillusJDR-2_IID_WP_015847338.1 282 KPVIYRVLDKTEGNRIADILIDGKTYTGREVREKLGKLSADFEWK..LDGDEIDITVEGFGHGVGMSOYGANGMAKSKSAEQIVYTYTGTIKVEQASKLAAKF....R

```

SpolIM – representative Bacilli

```

Bsub168_IIM_NP_390234.1      1 MRKI...SYKDMFLRHVKDHLSTYIFVSVLFFMGVIFGAIIVNSMTISOKEDLYYLSQFFGQLSDGKQAS.S.ADMFQOSIFHNAKYLGLMWILGIVTIGMPITFIMIFLFGK
Blicheniformis_IIM_AAU24024.1 1 MRKP...TIKELIFQHMKDHLSTYLFVSVLFLMGVIFGAVIVNSMTIGOKEDLYYLNQFFGQLSEGKAAS.S.KEMFLQSFHLNMMKYLGLMWILGIVTIGMPITFIMIFLFGK
Bhalodurans_IIM_BAB05245.1 1 MRKYGKKYGLKESMHRHLYDNQAIYLFIVLFLVGVIFGAVIVNSLSTOKHDLVYLYQFFGQVSDGHVAT.S.TEMFKQSFSHYAKYFGLMWILGIVTIGMPITFIMIFLFGK
Bclausii_IIM_BAD64314.1      1 MNNPNRVY...EAIARHVDNRRSIIYFIVFLIIGVIFGAVIVNSLSORHDLFTYLSQFFNEMKNGSVVT.SPSELVHVSFTHYGKYIIGLMWILGIVTIGMPITFIMIFLFGK
Bcereus_IIM_KFL78055.1      1 MWRK...NWQDRVMSHIQENSSIIYFNAVLFLMGVIFGAILVNSLQINOKQDLFFYLQRFQVQVSKGEFAI.A.GEMFRESEYFSOLKYLIGFIWILGIVTIGMPITFIMIFLFGK
Banthracis_IIM_YP_030252.1  1 MWRK...TWQDRVMSHIQENSSIIYFNAVLFLMGVIFGAILVNSLQINOKQDLFFYLQRFQVQVSKGEFAI.A.GEMFRESEYFSOLKYLIGFIWILGIVTIGMPITFIMIFLFGK
Oihyensis_IIM_BAC13806.1    1 MHKS...YVFNHFKHHATIIYFETIILFLTIGVIFGAVIVNSMDVVOKQDLFFYLERFFIQTTEEGSTLNN.NDILWQSFYHYIKYLGLMFIIFALSVTIGMPITFIMIFLFGK
Gkaustophilus_IIM_BAD76603.1 1 MRTH...PLKSAIAVHWRHSHASIIYFVIVFLFLMGVIFGAVIVNSLGFQKQDLFFYLLQFFGQVSKDNIAS.A.HDMFRQSYMHNKYLALMWWLGIIVTIGMPITFIMIFLFGK
Aflavithermus_IIM_ACJ33368.1 1 MRKR...SWTTTGWQHVQEHYTIYFVIVFLVFMGIIFGAVIVNSLFSOKEDLYYLLNFFGQVSAAGNMAN.A.HDVWKESEFLQNLKYLGLMWILAMSLVGLPITLILFLFGK
Bbrevis_IIM_BAH43355.1      1 MRSR...VG.QTIQSYAKHQSIYFIVLFTMGIIFGAVLVNSLPLSQKQDLVYGLQYFFNSLSDGDIPE.T.SAHFQOAFGHYAKIITAIMWWLGIIVTIGMPITFIMIFLFGK
PaenibacillusJDR2_IIM_WP_015843927.1 1 MRSP...AFQ...DAVKNQLTIYFVAVLFLVVGAIIFGLMVALTLEQQDLAADVNYIRLVNDGVTPD.A.AQSEWDSFIHFGKMLLWFLGLISVVGIPFVLAIDFLFGK

```

```

Bsub168_IIM_NP_390234.1      109 IVVGFVTCFLVNGMGVSCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Blicheniformis_IIM_AAU24024.1 109 IVVGFVTCFLVNGMGVSCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Bhalodurans_IIM_BAB05245.1 113 VVIGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Bclausii_IIM_BAD64314.1      111 LVIGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Bcereus_IIM_KFL78055.1      109 VVVGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Banthracis_IIM_YP_030252.1  109 VVVGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Oihyensis_IIM_BAC13806.1    107 VVVGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Gkaustophilus_IIM_BAD76603.1 109 VVVGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Aflavithermus_IIM_ACJ33368.1 109 LVVGFVTCFLVNSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Bbrevis_IIM_BAH43355.1      108 VVVGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
PaenibacillusJDR2_IIM_WP_015843927.1 105 ALVGFVTCFLVNSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...

```

SpolIM – representative Clostridia

```

Bsub168_IIM_NP_390234.1      1 MRKI...SYKDMFLRHVKDHLSTYIFVSVLFFMGVIFGAIIVNSMTISOKEDLYYLSQFFGQLSDGKQAS.S.ADMFQOSIFHNAKYLGLMWILGIVTIGMPITFIMIFLFGK
Blicheniformis_IIM_AAU24024.1 1 MRKP...TIKELIFQHMKDHLSTYLFVSVLFLMGVIFGAVIVNSMTIGOKEDLYYLNQFFGQLSEGKAAS.S.KEMFLQSFHLNMMKYLGLMWILGIVTIGMPITFIMIFLFGK
Bhalodurans_IIM_BAB05245.1 1 MRKYGKKYGLKESMHRHLYDNQAIYLFIVLFLVGVIFGAVIVNSLSTOKHDLVYLYQFFGQVSDGHVAT.S.TEMFKQSFSHYAKYFGLMWILGIVTIGMPITFIMIFLFGK
Bclausii_IIM_BAD64314.1      1 MNNPNRVY...EAIARHVDNRRSIIYFIVFLIIGVIFGAVIVNSLSORHDLFTYLSQFFNEMKNGSVVT.SPSELVHVSFTHYGKYIIGLMWILGIVTIGMPITFIMIFLFGK
Bcereus_IIM_KFL78055.1      1 MWRK...NWQDRVMSHIQENSSIIYFNAVLFLMGVIFGAILVNSLQINOKQDLFFYLQRFQVQVSKGEFAI.A.GEMFRESEYFSOLKYLIGFIWILGIVTIGMPITFIMIFLFGK
Banthracis_IIM_YP_030252.1  1 MWRK...TWQDRVMSHIQENSSIIYFNAVLFLMGVIFGAILVNSLQINOKQDLFFYLQRFQVQVSKGEFAI.A.GEMFRESEYFSOLKYLIGFIWILGIVTIGMPITFIMIFLFGK
Oihyensis_IIM_BAC13806.1    1 MHKS...YVFNHFKHHATIIYFETIILFLTIGVIFGAVIVNSMDVVOKQDLFFYLERFFIQTTEEGSTLNN.NDILWQSFYHYIKYLGLMFIIFALSVTIGMPITFIMIFLFGK
Gkaustophilus_IIM_BAD76603.1 1 MRTH...PLKSAIAVHWRHSHASIIYFVIVFLFLMGVIFGAVIVNSLGFQKQDLFFYLLQFFGQVSKDNIAS.A.HDMFRQSYMHNKYLALMWWLGIIVTIGMPITFIMIFLFGK
Aflavithermus_IIM_ACJ33368.1 1 MRKR...SWTTTGWQHVQEHYTIYFVIVFLVFMGIIFGAVIVNSLFSOKEDLYYLLNFFGQVSAAGNMAN.A.HDVWKESEFLQNLKYLGLMWILAMSLVGLPITLILFLFGK
Bbrevis_IIM_BAH43355.1      1 MRSR...VG.QTIQSYAKHQSIYFIVLFTMGIIFGAVLVNSLPLSQKQDLVYGLQYFFNSLSDGDIPE.T.SAHFQOAFGHYAKIITAIMWWLGIIVTIGMPITFIMIFLFGK
PaenibacillusJDR2_IIM_WP_015843927.1 1 MRSP...AFQ...DAVKNQLTIYFVAVLFLVVGAIIFGLMVALTLEQQDLAADVNYIRLVNDGVTPD.A.AQSEWDSFIHFGKMLLWFLGLISVVGIPFVLAIDFLFGK

```

```

Bsub168_IIM_NP_390234.1      109 IVVGFVTCFLVNGMGVSCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Blicheniformis_IIM_AAU24024.1 109 IVVGFVTCFLVNGMGVSCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Bhalodurans_IIM_BAB05245.1 113 VVIGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Bclausii_IIM_BAD64314.1      111 LVIGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Bcereus_IIM_KFL78055.1      109 VVVGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Banthracis_IIM_YP_030252.1  109 VVVGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Oihyensis_IIM_BAC13806.1    107 VVVGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Gkaustophilus_IIM_BAD76603.1 109 VVVGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Aflavithermus_IIM_ACJ33368.1 109 LVVGFVTCFLVNSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
Bbrevis_IIM_BAH43355.1      108 VVVGFTVCFVLSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...
PaenibacillusJDR2_IIM_WP_015843927.1 105 ALVGFVTCFLVNSQMGWNCFFLLSFSVSVLQNVLLIPAYLIMGTCAIATSLKLRIRQLFVKRSLHDAPIQWFGRYAFVLLVILFLALISLFFAYISPVLMKELTSRL...F...

```


49 **Fig. S1 – Multiple sequence alignment of DMP in representative endosporeformers**

50 Protein sequences of SpoIIP, SpoIID and SpoIIM from *B. subtilis* and *C. difficile* were aligned

51 with the sequences of retrieved orthologues in the representative Bacilli and Clostridia [1],

52 respectively, using T-Coffee [3]. Motifs identified after full HMM searches in all Bacilli and

53 Clostridia as well as the proposed catalytic residues are labelled: SpoIIP – motif 1, light

54 purple; motif 2, dark purple; SpoIID – motif 1, red; motif 2, dark orange; motif 3, light

55 orange. Catalytic residues in SpoIID and SpoIIP are shaded orange. Strictly conserved

56 residues are boxed in a white background, residues with similarity score above 0.7 (ESPrpt)

57 are shaded black. Alignment figures were generated with ESPrpt 3.0 [12].

58

- 59 [1] M.J.L. de Hoon, P. Eichenberger, D. Vitkup, Hierarchical evolution of the bacterial sporulation network,
60 Curr. Biol. 20 (2010) R735–45. doi:10.1016/j.cub.2010.06.031.
- 61 [2] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol.
62 215 (1990) 403–410. doi:10.1016/S0022-2836(05)80360-2.
- 63 [3] C. Notredame, D.G. Higgins, J. Heringa, T-Coffee: A Novel Method for Fast and Accurate Multiple
64 Sequence Alignment, J. Mol. Biol. 302 (2000) 205–217. doi:10.1006/jmbi.2000.4042.
- 65 [4] S.R. Eddy, Profile hidden Markov models, Bioinformatics. 14 (1998) 755–763.
66 doi:10.1093/bioinformatics/14.9.755.
- 67 [5] L.S. Johnson, S.R. Eddy, E. Portugaly, Hidden Markov model speed heuristic and iterative HMM search
68 procedure, BMC Bioinformatics. 11 (2010) 431. doi:10.1186/1471-2105-11-431.
- 69 [6] R.D. Finn, J. Clements, W. Arndt, B.L. Miller, T.J. Wheeler, F. Schreiber, A. Bateman, S.R. Eddy, HMMER
70 web server: 2015 update, Nucleic Acids Res. 43 (2015) W30–W38. doi:10.1093/nar/gkv397.
- 71 [7] T.J. Wheeler, J. Clements, R.D. Finn, Skylign: a tool for creating informative, interactive logos
72 representing sequence alignments and profile hidden Markov models, BMC Bioinformatics. 15 (2014)
73 7. doi:10.1186/1471-2105-15-7.
- 74 [8] J.J. Almagro Armenteros, K.D. Tsirigos, C.K. Sønderby, T.N. Petersen, O. Winther, S. Brunak, G. von
75 Heijne, H. Nielsen, SignalP 5.0 improves signal peptide predictions using deep neural networks, Nat.
76 Biotechnol. 37 (2019) 420–423. doi:10.1038/s41587-019-0036-z.
- 77 [9] L. Käll, A. Krogh, E.L. Sonnhammer, A Combined Transmembrane Topology and Signal Peptide
78 Prediction Method, J. Mol. Biol. 338 (2004) 1027–1036. doi:10.1016/J.JMB.2004.03.016.
- 79 [10] A. Krogh, B. Larsson, G. von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology
80 with a hidden markov model: application to complete genomes, J. Mol. Biol. 305 (2001) 567–580.
81 doi:10.1006/JMBI.2000.4315.
- 82 [11] D.W.A. Buchan, D.T. Jones, The PSIPRED Protein Analysis Workbench: 20 years on, Nucleic Acids Res.
83 (2019). doi:10.1093/nar/gkz297.
- 84 [12] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server,
85 Nucleic Acids Res. 42 (2014) W320–W324. doi:10.1093/nar/gku316.

86

87 **Table S1 – DMP orthologues identification in representative endosporeformers**

	% identity to query					
	BsIID	CdIID	BsIIP	CdIIP	BsIIM	CdIIM
Bacilli						
<i>Bacillus subtilis subsp. subtilis str. 168</i>	-	35.7	-	29.4	-	
<i>Bacillus licheniformis ATCC 14580</i>	70.9	37.7	71.2		79.2	
<i>Bacillus halodurans C-125</i>	48.3	34.5	42.2		54.4	
<i>Bacillus clausii KSM-K16</i>	44.9	33.9	42.6		52.8	
<i>Bacillus cereus ATCC 10987</i>	38.5	42.7	45.8		51.9	
<i>Bacillus anthracis str. Sterne</i>	53.2	43.4	45.4		49.8	
<i>Lysinibacillus sphaericus C3-41</i>	38.5	34.9	30.0			
<i>Oceanobacillus iheyensis HTE831</i>	45.8	40.4	44.1		42.2	
<i>Geobacillus kaustophilus HTA426</i>	58.1	39.7	51.4	26.9	56.3	
<i>Anoxybacillus flavithermus WK1</i>	53.4	34.7	48.5	30.1	54.3	
<i>Brevibacillus brevis NBRC 100599</i>	46.5	32.6	37.0	27.1	48.0	
<i>Paenibacillus sp. JDR-2</i>	47.4	39.3	33.9	28.2	38.7	
Clostridia						
<i>Caldanaerobacter subterraneus subsp. tengcongensis MB4</i>	38.8	43.5	25.6	34.7	31.0	
<i>Moorella thermoacetica ATCC 39073</i>	44.0	37.7	31.3	29.6	39.6	
<i>Carboxydotherrmus hydrogenoformans Z-2901</i>	35.6	38.1		29.6	34.2	
<i>Desulfitobacterium hafniense Y51</i>	42.0	29.8	30.7	33.7	28.7	
<i>Hungateillostridium thermocellum ATCC 27405</i>	33.4	41.4		32.1	30.9	
<i>Desulfotomaculum acetoxidans DSM 771</i>	36.1	34.6	27.9	25.4	33.8	
<i>Clostridium acetobutylicum ATCC 824</i>	36.4	44.1	24.5	29.5	31.1	29.9
<i>Clostridium botulinum A str. ATCC 3502</i>	35.1	37.6	26.0	33.8	28.9	31.5
<i>Clostridium perfringens str. 13</i>	40.8	41.8	30.0	29.5	23.8	
<i>Clostridium difficile 630</i>	35.7	-	29.4	-		-
<i>Alkaliphilus oremlandii OhILAs</i>	37.9	40.8	33.2	37.9	31.2	
<i>Symbiobacterium thermophilum IAM 14863</i>	35.3	33.5			26.7	

88