Supplemental Materials and methods

Cloning and expression of the diffocin gene clusters in *Bacillus subtilis*

pETcoco1 (Novagen) was the starting backbone vector that was first modified to remove the two XhoI sites with primers AV1419-AV1420 which have BbsI ends (primer sequences are shown in Supplemental table 2). Then a region was amplified from pETcoco1 DNA with these primers, and the PCR product was cut with BbsI and was ligated back into the larger pETcoco1 vector fragment that was previously cut with XhoI. The result of this ligation destroyed the two XhoI sites. This plasmid was the further modified by a similar strategy to destroy the EcoRI sites using primers AV1416 and AV1245. The resulting vector is termed pSW251.

Vector pUC19 was used to accept fragments of the diffocin cluster to assemble into pETcoco. The polylinker of pUC19 was modified by cutting with EcoRI and HindIII and ligating in oligos AV1372, AV1373, AV1374, AV1375. This changed the polylinker to NotI-NheI-KpnI-XhoI-EcoRV-BstBI-BbsI-EcoRI-NsiI-SphI-BamHI-AscI. The construct was termed pSW232. Three fragments of the diffocin cluster were individually amplified by PCR from *C. difficile* 16 (CD16) DNA and cloned into pSW232. The 5' fragment was amplified with primers AV1368 and AV1289 which have NotI and XhoI ends respectively. The middle fragment was amplified with primers AV1288 and AV1366 which have XhoI and EcoRI ends respectively. The 3' fragment was amplified with primers AV1367 and AV1300 which have EcoRI and BamHI ends respectively. These three PCR fragments were separately cloned into pSW232, and termed pSW243, pSW241, and pSW242 for the 5', middle, and 3' portions respectively.

The three diffocin cluster fragments were excised from the pSW241-3 vectors. The pSW243 was cut with NotI-XhoI, pSW241 was cut with XhoI-EcoRI, and pSW242 was cut with EcoRI- AscI (note that this AscI sites is part of the modified SW232 polylinker). These three fragments were

assembled into pSW251 that had been cut with NotI and AscI. The resulting plasmid was termed pDG461 and contains the entire CD16 diffocin cluster.

To make a *B. subtilis* integration vector we used sequences from pDR111 (gift from David Rudner, Harvard Medical School) that includes portions of the *amyE* gene flanking a cloning/promoter region and a spectinomycin marker. This plasmid was modified by cutting with NcoI and SpeI and religating to create pDG481. The pDG481 polylinker was then modified by cutting the vector with HindIII and SphI and ligating in oligos DG1 and DG2. This added NotI and AscI sites to create pDG482. The region containing the entire *amyE* front and back region with the modified polylinker was then amplified using primers DG9 and DG10 both of which have BsaI ends. This fragment was cloned into the NotI and AscI sites of pSW251 (resulting in destruction of the two sites) creating pDG487. Note that there are new NotI and AscI sites that were introduced by the modified poly linker insertion. The NotI/AscI fragment containing the diffocin cluster from pDG461 was then excised and cloned into the NotI/AscI site of pDG487. This new construct was termed pDG488 and is the vector used to introduce the entire diffocin gene cluster into *B. subtilis* BDR123 (see below) to generate BDR123-488. To delete ORFs 1377-1379, a backbone was made by cutting pDG488 with Sph1 which has sites outside the integration sequence and one within 1374. Two fragments were amplified from pDG488. The first used primers DG13/14 which generated a fragment from the SphI site to the end of ORF 1376 in which an AscI site was incorporated. The second fragment, also with SphI/AscI sites used primers DG15/16 which span from the SphI site outside the integration site to sequence downstream of 1379 (AscI incorporated in the primer). The three way ligation resulted in pDG491. This was transformed into BDR123 to generate BDR123-491.

Diffocin 4 was cloned similarly to diffocin 16 with some modifications due to the absence of the EcoRI site within the CD4 gene cluster. The pSW251 was modified to have an XhoI site in the polylinker using oligos DG211 and DG212 (NotI and AscI). This created vector pDG577. The diffocin cluster was amplified in three fragments. The first was using primers DG210 and AV1288 (XhoI/NcoI). The second used primers DG209 and DG15 (NcoI and AscI). These two were cloned into pDG577 (cut with XhoI/AscI) to create pDG578. The third fragment was amplified using AV1368 and AV1289 (XhoI/NotI) and cloned into pDG578 (cut with XhoI/NotI) to create pDG579. The integration vector was made by taking the NotI/AscI fragment from pDG579 and cloning it into pDG487. This construct was termed pDG580 and was transformed into BDR123 to generate BDG12.

Deletion of the PBSX locus.

The PBSX prophage knockout was constructed by following the procedure outlined in Liu S, et al. 2008. Microbiology. 154:2562-2570. Briefly, the promoter of the *araA* gene, PCR amplified from *B. subtilis* strain BDR11 (David Rudner, Harvard Medical School), was fused to a neomycin resistance gene (*neo*), PCR amplified from plasmid pUB110, by overlapped extension PCR. Overlapped extension PCR was then used to fuse this arabinose regulated neomycin gene to about 1 kb of sequence flanking each end of the *araR* gene, PCR amplified from BDR11. This PCR product was transformed into BDR11 to make strain BDG2, deleting the *araR* gene and replacing it with a constitutively expressed *neo*. The deletion of *araR* and replacement with P_{araA}-neo was confirmed by PCR and by the conferral of resistance to kanamycin. This strain is functionally equivalent to AR1 from Lui et al., 2008 (17).

Next, a construct was made to knockout the PBSX locus itself. To make this construct, the following five PCR products were spliced by overlapped extension PCR in order, into one large

product: 1kb of sequence 5' of the *xlyB* gene, amplified from BDR11; 1 kb of sequence 3' of the *xlyA* gene, amplified from BDR11; a chloramphenicol resistance gene, *cat*, amplified from plasmid pJW034; *araR,* amplified from BDR11; and finally, 1 kb of sequence containing the *xlyB* gene, amplified from BDR11. The overlapped extension PCR product was cloned into the XmaI and SpeI sites of pBluescript SK+. This construct was then linearized with ScaI, transformed into strain BDG2, and plated onto LB agar plates supplemented with 5 µg/ml chloramphenicol. Colonies were picked from this plate and patched onto LB agar plates supplemented with either 5 µg/ml chloramphenicol, or 20 µg/ml kanamycin. Successful recombinants were those that are resistant to chloramphenicol, but sensitive to kanamycin due to repression of the *neo* gene by AraR. These strains will have replaced the PBSX target with the *cat* and *araR* genes. Such isolates were selected after growth for 4 hours in LB broth with no antibiotic selection, and then plated onto LB agar plates supplemented with 20 µg/ml kanamycin. This was to select for recombinants that loop out the *cat* and *araR* genes resulting in a clean deletion of the target PBSX genes. The colonies that grew on these plates were tested by colony PCR for the presence of a PBSX knockout. The deletion was confirmed by PCR in strain BDG9. To further optimize expression, additional constructs were made by first switching the selective marker from spectinomycin to chloramphenicol and by designing the integration construct as a single crossover into *amyE*. The pDG487 was modified by digestion with AscI and NheI to remove the *amyE* 5' portion and *lacI*, and cloning in oligos DG310 and DG 311 which destroys the NheI site and adds an RsrII site, generating pDG588. This was cut with AscI and NotI, and the AscI/NotI fragment from pDG579 containing the diffocin 4 cluster was ligated in to generate pDG589. This construct was transformed into BDG9 to generate BDG21. BDG 21 contained the

entire diffocin 4 locus as well as the spectinomycin marker flanked by identical back ends of the *amyE* gene.

To switch the selective marker, the following construct was made. The pDG481 was cut with BbsI and PsiI to remove the front portion of the a*my*E gene. Two PCR fragments were then prepared for a three way ligation into this vector; the first was from pJW034 (David Rudner, Harvard Medical School) using primers DG488 and DG489 that contains the *cat* gene. The second fragment was amplified from diffocin 4 using DNA primers DG367 and DG368. The three way ligation resulted in pDG621. This was transformed into BDG21 and recombinants resistant to chloramphenicol and sensitive to spectinomycin were selected resulting in strain BDG45.

Switching the bactericidal spectrum of diffocins.

To replace the diffocin 4 *1374* with that of diffocin 16, a backbone was made by cutting pDG580 with AscI and PmlI and isolating the larger fragment. Two other fragments were amplified from other constructs. The first was from plasmid pDG461 using primers DG281 (PmlI) and DG316 (BspQI). The second was amplified from pDG579 using primers DG284 (AscI) and DG317 (BspQI). The three way ligation created pDG587. This plasmid contained the diffocin 16 *1374* and part of ORF 1373. To reconstruct the diffocin 4 ORF 1373, such that the only replacement in the cluster is *1374* gene, the following PCR products were produced and cloned into the pDG461 PmlI/BspQI backbone: pDG587 was amplified using primers DG284 and DG 339 and cut with AscI and BsmBI and pDG579 was amplified with primers DG281 and DG340 and was cut with BsmBI and PmlI. This three way ligation resulted in pDG603. pDG603 was transformed into *B. subtilis* strain BDG9 to make BDG26 which was transformed again with pDG621 to produce BDG55.

Supplemental Table 1. Bacterial strains used in this study. All *C. difficile* strains are human clinical isolates.

Species	Strain designation	Ribotype	Source	Comments
Clostridium difficile	CD4	024	L-C Fourtier	Diffocin producer
Clostridium difficile	CD16	027	L-C Fourtier	Diffocin producer
Clostridium difficile	630	012	ATCC	
Clostridium difficile	43593	060	ATCC	Diffocin producer
Clostridium difficile	19135	001	Diane Citron	
Clostridium difficile	19137	015	Diane Citron	
Clostridium difficile	19099	002	Diane Citron	
Clostridium difficile	19145	153	Diane Citron	
Clostridium difficile	19103	001	Diane Citron	
Clostridium difficile	19155	NA	Diane Citron	
Clostridium difficile	19104	027	Diane Citron	
Clostridium difficile	19142	NA	Diane Citron	
Escherichia coli	Top10		Invitrogen	Electrocompetent cells for subcloning.
Bacillus subtilis	BDR11		David Rudner	Expression strain for diffocins
Bacillus subtlis	BDR123		David Rudner	Expression strain for diffocins
Clostridium acetobutylicum	824		ATCC	
Clostridium sordellii	9714		ATCC	
Clostridium sporogenes	3584		ATCC	
Clostridium biofermentans	638		ATCC	
Listeria Ivanovvii	19119		ATCC	
Listeria monocytogenes	23074		ATCC	
Listeria innocua	33090		ATCC	
Bifidobacterium breve	15700		ATCC	
Bacteroides fragilis	23745		ATCC	
Bacteroides fragilis	25285		ATCC	

Supplemental Table 2. Oligonucleotides/primers used in this study.

Primer name	Sequence 5'-3'
AV1245	ccccttgaagacccaatttcgtatggcaatgaaagacgg
AV1288	gaagaaagagatgggactcgagatg
AV1289	tatacatctcgagtcccatctctttc
AV1300	tcccccggatccacactacaatctactctaaactcagg
AV1367	ggaaaagggattgctaatagtg
AV1368	ttttttgcggccgcaatacccactacaccttcgtc
AV1372	aattgcggccgcagctcgctagcggtacctcgaggatatcttcgaagaagacacatccg
AV1373	aattccgggatgcatgcctctaggatccggcgcgcc
AV1374	agctggcgcgccggatcctagaggcatgcatcccggaattcggatgtgtcttctt
AV1375	cgaagatatcctcgaggtaccgctagcgagctgcggccgc
AV1416	ttccttgaagacctaatttggggcaatcccgcaaggag
AV1419	tttcttgaagaccatcgaagcaccaccaccaccaccactg
AV1420	ttttttgaagacaatcgaagggcttcgccctgtcgctcgac
DG1	ggcgcgccactagtaccggtgccatggcggccgc
DG2	agctgcggccgccatggcaccggtactagtggcgcgcccatg
DG9	ttccttggtctcacgcgaacaaaattctccagtcttc
DG10	ttccttggtctcaggccgtcgcgactaagaaaatgcc
DG13	gtgagcggataacaattccc
DG14	agattgtagtgtggatccgg
DG15	tccttcggcgcgcctcaaatttaagcttaactcc
DG16	tttagggactactcactcgc
DG209	actggataccatggtacttc
DG210	tgaagtaccatggtatccag
DG211	ggccgcctcgaggg
DG212	cgcgccctcgaggc
DG281	gtataagccacgtgctgaag
DG284	aggcgcgcctcaaatttaag
DG316	tcttctgctcttcaagggtctccaagatttaaatcttct
DG317	ctccttgctcttcacctatatatcaattaataaataaaac
DG339	tcctctcgtctcaatgaagcaaaataaacttttac
DG340	tcctctcgtctcttcattataaaaacctcctaattat

Supplemental Figure 1S. A) Clustal analysis of ORF 1374 from srtains CD4, CD16, CD43593, and 630. B) Dendrogram of the clustal analysis.

A.

```
4               MKRTKLLQRGNFFGDKNMVVDEFDEGYDNYDFINFFTGCCNYTFGLKNNNILYGCGDNSN
16              MKQNKLLQRGAYFNDKNILIDDFDKRYNDYDFVEFFTGISNSTFGLKSDGNLYACGDNTG
43593           MKQNKLLQRGAYFNDKNILIDDFDKRYNDYDFVEFFTGISNSTFGLKSDGNLYACGNNTG
630             MKQNKLLQRGAYFNDKNILIDDFDKRYNDYDFVEFFTGISNSTFGLKSDGNLYACGDNTG
                **:.****** :*.***:::*:**: *::***::**** .* *****.:. **.**:*:.

4               FQLGLGEDNTTRKLFTKIPNISTNIKKVACGESHAVILTSDGELLVAGINTDGQMGLGLE
16              FQLGLGKDSSERRMFSKVK--IDNVKYVSCGSKHSVAVTKDGFAYGAGTSNVGQLGVIES
43593           FPLGLGKDSSERRMFSKVK--IDNVKYVSCGSKHSVAVTKDGFAYGAGTSNVGQLGVIES
630             FQLGLGKDSSERRMFSKVK--IDNVKYVSCGSKHSVAVTKDGFAYGAGTSNVGQLGVIES
                * ****:*.: *::*:*:     *:* *:**..*:* :*.**    ** .. **:*:  .

4               KVGKTVSTFEKVPEIKGVKDIACGLQSTYLLYNDGTLYVAGNNLYGQLGLGTNGASANVN
16              TVYYEFTKLP----IDDVKTVACGYDFTFVLKNDGTLYSAGLNSSGQLGLG---DTNNRA
43593           TVYYEFTKLP----IDDVKTVACGYDFTFVLKNDGTLYSAGLNSSGQLGLG---DTNNRA
630             TVYYEFTKLP----IDDVKTVACGYDFTFVLKNDGTLYSAGLNSSGQLGLG---DTNNRV
                .*   .:.:     *..** :*** : *::* ****** ** *  ******    : *

4               TFTKVDVDNVKAVFSYNKSAFIIKNDNKCYSTGFNNQGQLGLGDKNNRDLFSLVS-INDV
16              TFTKVNIDSVKDVVTYNQSVFIIKMDGTAHACGLNSNGQLGINSTLNKSVFNKIEGMDNV
43593           TFTKVNIDSVKDVVTYNQSVFIIKMDGTAHACGLNSNGQLGINSTLNKSVFNKIEGMDNV
630             TFTKVNIDSVKDVVTYNQSVFIIKMDGTAHACGLNSNGQLGINSTLNKSVFNKIEGMDNV
                *****::*.** *.:**:*.**** *...:: *:*.:****:... *:.:*. :. :::*

4               KTIACGSEHTVLMTYNNDIYGCGK----EKCFGNALQSSLFTKIEEVNIKTIACGHGNTM
16              KQIACGSSHTILIKNDGTMYTTGYNGVGQLGTGNNNNSIVFTLSSINNVKYASCGNNHTM
43593           KQIACGSSHTILIKNDGTMYTTGYNGVGQLGTGNNNNSIVFTLSSINNVKYASCGNNHTM
630             KQIACGSSHTILIKNDGTMYTTGSNGYGQLGTGNNNNSIVFTLSSINNVKYASCGNNHTM
                * *****.**:*:. :. :*  *     :   **  :* :**  .  *:*  :**:.:**

4               LIDNKGTLKVAGNNDIYQLGIANYSENIDNSFIDLKNIVAKNIFIGLSHSILIDSNNDSY
16              ILKYDNTLFSTGQNNYGQLANANKDVASRNTFAKVNVENIKDIKCGSQFNFLINGSKEIF
43593           ILKYDNTLFSTGQNTYGQLANANKDVASRNTFAKVNVENIKDIKCGSQFNFLINGSKEIF
630             ILKYDNTLFSTGQNNYGQLANANKDVASRNTFVKVNVENIKDIKCGSQFNFLINGSKEIF
                ::. ..**  :*:*   **. ** .    *:* .::    *:*  * ...:**:..:: :

4               CTGDNTYGQLGSFFDDMHIVEFKKMDSEKYSYSNYINLIKSEDKLTLLKEEMEIKDIELP
16              VSGCNLAGQLGSFFHTTFLYEFSNVQSS--NLDNYSGLLVNDDYLYVTKDNSEFLNVKLS
43593           VSGCNLAGQLGSFFHTTFLYEFSKVQSS--NLDNYSGLLVNDDYLYVTKDNSEFLNVKLS
630             VSGCNLAGQLGSFFHTTFLYEFSKVQSS--NLDNYSGLLVNDDYLYVTKDNSEFLNVKLS
                 :* *  *******.  .: **.:::*.  . .** .*: .:* * : *:: *: :::*.

4               LDIHSVRDVVFSPYCTLVILGNGDVYGLGNNRYKGMGSDLPSQLNELTKLSISNVKSIVA
16              DNFQDYKKIELTDSNMFIVMNDGTLYACGLNNYGQLGLGDTVNRSVMTKVDIDNVLDIKG
43593           DNFQDYKKIELTDNNMFIVMNDGTLYACGLNNYGQLGLGDTVNRSVMTKVDIDNVLDIKG
630             DNFQDYKKIELTDNNMFIVMNDGTLYACGLNNYGQLGLGDTVNRSVMTKVDIDNVLDIKG
                 :::. :.: ::    ::::.:* :*. * *.*  :* . . : . :**:.*.** .* .

4               SKNISGGIFYIKNDDTCYYSGPNSNSIAGVLP-SNSDVFKKISIDNVKKVVINTDLSNWF
16              NGNST---FVLKNNGTLYSCGLNSNGQLGLRDEVNRNIFTKIEIENVKDFCVGSNY----
43593           NGNST---FVLKNNGTLYSCGYNSSGILGLKDNTNRNIFTKIEIENVKAFCVESNY----
630             NGNST---FVLKNNGTLYSCGYNSSGILGLKDNTNRNIFTKIEIENIKEFCVESNY----
                . * :   * :**:.* * .* **..  *:    * ::*.**.*:*:* . : ::

4               SLIVTNNKQIYTSGKSSSYVNGLSNALISQYTEISLSNVTDAYSSYNATFIVVDEKKVYA
16              --VIALNHSKEVYGWGNNPYNNIEKTSNYPYKQG-ISNIEKIAAYDYSVYMINSEGKLYV
```

```
43593          --IVVLNHSKELYGWGNESYIVYGNSRNYPYKDTRVSNVEKIATWSDTLYILDSTGATKT
630            --IVALNHSKELYGWGNQSYIVYGDNRNYPYKDTRVSNVEKIATWSDTLYILDSTGATKT
                 ::. *:.    * ...      .     *.:  :**: .  :   : ::: .     .

4              TGINTNYLLGFSTSDGSNVNLGLLSDWYYINISGSSYSRVSCTNNITKINNIIIYEYVTV
16             SGYNYNYQLGKGNNSNQSK---ALVSQCRTNSTSSTSNGLR--TLPKITNVFPFYDGCAI
43593          IGYSYNGSGGYPAPSSSST---YRDGGYINKNTSYRTLEFYNTSKTKLVNLFAFYNGCVF
630            IGYSYNGSGGYPAPSSSST---YREGGYINKNTSYRTLEFYNTSKTKLVNLFAFYNGCVF
                * . *   *    ....       .    : :.     .   .  .  * : :*:  ..

4              FCTNIGSFLTGYHGTSWTKPTDSSYRVQYQGISYAGYLDSYIYNYYPTRCTQSSSSTTFA
16             IDEGGYVYLTGYHGYLRTLNSSPSISDYSR---YGTFIEATNSNHN----------TYFI
43593          VDENGLAYCIGENNINFRGNSTTNENNSLR---FINNSGVYYTNTDGTDYTCYQWTYKLI
630            VDENGLAYCIGENNINFRGGSTTNENNSLR---FINNSGVYYTNTDGTDYTCYQWTYKLI
               .  .   :  * :.      : ..     :   :         *             :

4              YLYNGESSSNLKNVNPDNLLISGGSSYIHQYGRNYLNNQSSNNIAASNINSGPITSDKAI
16             QETDFSGIEKVIGMSNNILFFKKGSSYITGYP-KTFGSTITGHRSYTSINSESSNLGSN-
43593          RCSIFDSPQNIIGNSKNILYLSKNNSTFKCTG-NCITYGINSQNWYSYFSDSSNGAIALG
630            RCSIFDSPQNIIGNSKNILYLSKNNSTFKCTG-NCITYGINSQNWYSYFSDSSNGAIALG
                    .. .:: . . : * :. ..* :     : :    ..:   : :.. .

4              FLYKALLYLSSNTLYGFGNISESAKELDVSDTQDGYNATNYKKVMKNIKNIFIPPYDLSR
16             ----FIIYHSNSKLYGKG----IANSGQFGNSTNIDGTSNYDTGLKDIKDIIVK------
43593          NEFILKNYSGECLLKGYG----KATNGEFGNSTNISSISNYDTGLKDIKDIIVK------
630            NEFILKNYSGECLLKGYG----KATNGEFGNSTNISSISNYDTGLKDIKDIIVK------
                      * ..  * * *     *.. :..:: :  . :**.. :*:**:*::

4              DKTRFAILTDKSLFICGYNSKGTHGISVNSSLNLNNKINYNKKNSSSEISSNIQEIYSHS
16             GNTVVVVDKNNNIYVTGMNQNNKLGIGEYNNEPVKKFTNITEQSNSFIFMDDIKEITTSR
43593          NNTVVVVDKNNNIYVTGANQFNKLGIGEYNNQPIKKFTNITEQSNSFIFMDDIKEITTSR
630            NNTVVVVDKNNNIYVTGANQFNKLGIGEYNNQPIRKFTNITEQSNSFIFMDDIKEITTSR
               .:* ..: .::.::: * *. .. **.  ..  :.:  * .::..*  : .:*:** :

4              KSTYLLTNNNMLYSVGLNDVGQLGVGDEINRKVFTKINIDNIKSINVNRFTDNSKHAFAI
16             NTMFIVKNDGTAYATGNNSSGQLGLGDTINRNKFTQINLDNIKKIST---SIDGNTTFAI
43593          NTMFIVKNDGTAYATGNNSSGQLGLGDTINRNKFTQINLDNIKKIST---SIDGNTTFAI
630            NTMFIVKNDGTAYATGNNSSGQLGLGDTINRNKFTQINLDNIKKIST---SIDGNTTFAI
               :: :::.*:.  *:.* *. ****:** ***: **:**:****.*..   : :.: :***

4              KNDNTCYAVGLNNSGQLGIGDNVNRNIFTKINVENVKYVAVYGNTSLLLTNDGLLYGAGN
16             RNDGTLYSTGLNTKGQLGLGDIVNRNTFTKVNIQNVRDVVLGTTHSHAIKDDNTLYSCGE
43593          RNDGTLYSTGLNTKGQLGLGDIVNRNTFTKVNIQNVRDVVLGTTHSHAIKDDNTLYSCGE
630            RNDGTLYSTGLNTKGQLGLGDIVNRNTFTKVNIQNVRDVVLGTTHSHAIKDDNTLYSCGE
               :**.* *:.***..****:** **** ***:*::**: *.:  . *  :.:*. **..*:

4              NGKGQLGLGDTTSR---NIFTRIPINGVRDVYLCNDVSIIVKNDNTCYVCGLVNGYFGFT
16             NTHGQLGLGSESNHPDVLTFTVNNITNVRDVYCSDTTTFIVKDTNIAYCCGYNNNSQLGM
43593          NTHGQLGLGSESNHPDVLTFTVNNITNVRDVYCSDTTTFIVKDTNIAYCCGYNNNSQLGM
630            NTHGQLGLGSESNHPDVLTFTVNNITNVRDVYCSDTTTFIVKDTNIAYCCGYNNNSQLGM
               * :******. :.:     **   *..***** .: .::***: * .* **  *.

4              EGSISTFTKINIENVKSVVTAGSEATFFITNDNMIYTTGKKERVFFSTETNDIKGIRVIN
16             GNTTDQYSFIKCMENVKEVIPNEINTYIITIYNTAYSTGLNTDYCLGLNSNSNQSSFSEI
43593          GNTTDQYSFIKCMENVKEVIPNEINTYIITIYNTAYSTGLNTDYCLGLNSNSNQSSFSEI
630            GNTTDQYSFIKCMENVKEVIPNEINTYIITIYNTAYSTGLNTDYCLGLNSNSNQSSFSEI
                .: . :: *:  :  . * ...  *::**  *  *:** :    :. ::*. :.

4              NIINAKKIVVN-GYTSAILTNDNKLFVG----GLSGYGSIANNNNTNSVEDVKDVFVTAN
16             PISNVVKVAPNRNNAVLLLTSEGDVYTAGKCSNGSGTGSETPEKIKKIASKAKDIGMNYR
43593          PISNVVKVAPNRNNAVLLLTSEGDVYTAGKCSNGSGTGSETPEKIKKIASKAKDIGMNYR
630            PISNVVKVAPNRNNAVLLLTSEGDVYTAGKCSNGSGTGSETPEKIKKIASKAKDIGMNYR
                * *. *:. * . :  :**.:..::..    . ** ** : :: .: ....**: :. .
```

```
4               NTLYIDNNNNLISSGRDTYGISDESYRDMSVPYYKVSIKKDVDTVFSSYNTIFIKDIYGK
16              CGHYVSDNGDLYGTGFNNNGQLGVGDVTKRDTFIKTNTR-VKKILPLEYANIAIKDTNDI
43593           CGHYVSDNGDLYGTGFNDCGQLGVGDVTKRDTFIKTNTR-VKKILPLEYANIAIKDTNDI
630             CGHYVSDNGDLYGTGFNDCGQLGVGNVTKRDTFIKTNTR-VKKILPLEYANIAIKDTNDI
                   *:.:*.:* .:* :  *  . .      .: *.. :   . :  .* .* ***  .

4               FYSSTRDNRYNHLGIHHRYDNDKNEALEGSLHSYFKTDNTSDKIVFNKKNEKLVMFNDKY
16              YICGLNN--YGQLGVGNRYDSRNN---DNRIFNYKHMNFVMGDLTSIKNRHNFILLNNKI
43593           YICGLNN--YGQLGVGNRYDSRNN---DNRIFNYKHMNFVMGDLTSIKNRHNFILLNNKI
630             YICGLNN--YGQLGVGNRYDSRNN---DNRIFNYKHMNFVMGDLTSIKNRHNFILLNNKI
                : .. .:  *.:**: :***. :*   :. :..* : : . ..:.  *:..::::::*:*

4               IKTNNKYINYKNIFKDNFK--YTSIILPFEVSDIDISKTHSLAVAKDGKLYGIGSNSYKE
16              VIPTTKDIDYGLVLGNLYKGDLYTELPYEDIKEVSISKTHIIILLNDGTMYGCGTNYHGE
43593           VIPTTKDIDYGLVLGNLYKGDLYTELPYEDIKEVSISKTHIIILLNDGTMYGCGTNYHGE
630             VIPTTKDIDYGLVLGNLYKGDLYTELPYEDIKEVSISKTHIIILLNDGTMYGCGTNYHGE
                : ...* *:*  :: : :*    : :   ::.::.***** : : :**.:** *:* : *

4               INQTLEDIELLTLTEVNISDVKKVACGDNYSYIIKTDNTLWSYGKNTEYQLGVGHNNDVR
16              LLQDLSINQVDEFVQINVSDVKHVSCGDNFTYFIKSDDSLWSIGKNSEYQLGIGHNNPVT
43593           LLQDLSINQVDEFVQINVSDVKHVSCGDNFTYFIKSDDSLWSIGKNSEYQLGIGHNNPVT
630             LLQDLSINQVDEFVQINVSDVKHVSCGDNFTYFIKSDDSLWSIGKNSEYQLGIGHNNPVT
                : * *.  ::  :.::*:****:*:****::*:**:*::*** ***:*****:**** *

4               ELQKVTGLPSVKDISIYNSMTLVLTNEGELYAQGYNTNGLFGLGESEKDKIIRTFTKVLT
16              ELQRITTISSCKEVHCGKNYTLVVTTGNELFVQGYNDKGALGLGSDSENTIIKFFTKALT
43593           ELQRITTISSCKEVHCGKNYTLVVTTGNELFVQGYNDKGALGLGSDSENTIIKFFTKALT
630             ELQRITTISSCKEVHCGKNYTLVVTTSNELFVQGYNDKGALGLGSDSENTIIKFFTKALT
                ***::* :.* *::   :. ***:*. .**:.**** :* :***...::.**: ***.**

4               NVKEIKSHNDDHILVIKNDNSLWITGKNKSMYKISISITDLYEFTKIPIPEHLNDILDIE
16              DIREIKSYGSDHILVLKNDNSVWVTGKNRDVYKIEQPVEFLKEFTIVPISEDVNTVKDVL
43593           DIREIKSYGSDHILVLKNDNSVWVTGKNRDVYKIEQPVEFLKEFTIIPISEDVNTVKDVL
630             DIREIKSYGSDHILVLKNDNSVWVTGKNRDVYKIEQPVEFLKEFTIVPISEDVNTVKDVL
                :::****:..*****:*****:*:****:.:***. .:  * *** :**.*.:* : *:

4               LSDDTIYMITKVDTSKASIEIVEKSISQVRVVVQDPNNVIEKLEMFINDELISTKTNLEI
16              ATDNTLYIISEVGTTNAAIEITEKSISSIKIKIQDPNKDISRIEMLINGESVKSVSDLTT
43593           ATDNTLYIISEVGTTNAAIEITEKSISSIKIKIQDPNKDISRIEMLINGESVKSVSDLIT
630             ATDNTLYIISEVGTTNAAIEITEKSISSIKIKIQDPNKDISRIEMLINGESVKSVSDLIT
                 :*:*:*:*::*.*::*:***.*****.::: :****: *.::**:**.* :.: ::*

4               NSIIFEIPQNKIVLGENKILIKASSPTGDLYSSMFIFKSETGLKVKKDSILMINNKVYSI
16              EKISFEVPPDKIKIGENKILFRAYCKGDDLYASLFIFKESTGNSIIKDSYVMIGNRMYKV
43593           EKISFEVPPDKIKIGENKILFRAYCKGDDLYASLFIFKESTGNSIIKDSYVMIGNRMYKV
630             EKISFEVPPDKIKIGENKILFRAYCKGDDLYASLFIFKESTGNSIIKDSYVMIGNRMYKV
                :.* **:* :** :******::* .  .***:*:****..** .: *** :**.*::*.:

4               INITENNTDLIVTLNEGLKDDMMENNPIYQLINKTKVQVKINKSDLFKDMKLVEIKKSDS
16              VNTTSNEQDITITLDRGLEEDLNLGDPIYQLINKTKVQVKINKSDLFKDMKLVEIKKSDS
43593           VNTTSNEQDITITLDRGLEEDLNLGDPIYQLINKTKVQVKINKSDLFKDMKLVEIKKSDS
630             VNTTSNEQDITITLDRGLEEDLNLGDPIYQLINKTKVQVKINKSDLFKDMKLVEIKKSDS
                :* *.*: *: :**:.**::*:  .:**********************************

4               SYQEIYELEEANIKSAQPKIIVEKGDKWTAIKRPSMIFRYDAENNEPQA
16              SYQEIYELEEANIKSAQPKIIVEKGDKWTAIKRPSMIFRYDAENNEPQA
43593           SYQEIYELEEANIKSAQPKIIVEKGDKWTAIKRPSMIFRYDAENNEPQA
630             SYQEIYELEEANIKSAQPKIIVEKGDKWTAIKRPSMIFRYDAENNEPQA
                **********************************************
```

B.