Sequence alignment for phytases *Aspergillus ficuum, Aspergillus niger* and *E.coli* phytases.

1 *Aspergillus ficuum* phytase

A. ficuum phytase accession code Q8J255
(http://www.uniprot.org/uniprot/Q8J255)

LAVPASRNQSTCDTVDQGYQCFSETSHLWGQYAPFFSLANKSAISPDVPAGCQVTFAQVL

SRHGARYPTDSKGKKYSALIEEIQQNATTFEEKYAFLKTYNYSLGADDLTPFGEQELVNS

GVKFYQRYESLTRNIVPFIRSSGSSRVIASGNKFIEGFQSTKLKDPRAQPGQSSPKIDVV

ISEASTSNNTLDPGTCTVFEDSELADDIEANFTATFVPSIRQRLENDLSGVTLTDTEVTY

LMDMCSFDTISTSTVDTKLSPFCDLFTHEEWINYDYLQSLNKYYGHGAGNPLGPTQGVGY

ANELIARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFSHDNGIISILFALGLYNGTK

PLSSTTAENITQTDGFSSAWTVPFASRMYVEMMQCQSEQEPLVRVLVNDRVVPLHGCPVD

ALGRCTRDSFVKGLSFARSGGDWAECFA

2 *Aspergillus niger* phytase

A. niger phytase accession code A2Q1G7
(http://www.uniprot.org/uniprot/A2QIG7)

VPASRNQSTCDTVDQGYQCFSETSHLWGQYAPFFSLANKSAISPDVPAGCHVTFAQVLSR

HGARYPTDSKGKKYSALIEEIQQNATTFEGKYAFLKTYNYSLGADDLTPFGEQELVNSGV

KFYQRYESLTRNIVPFIRSSGSSRVIASGNKFIEGFQSTKLKDPRAQPGQSSPKIDVVIS

EASTSNNTLDPGTCTVFEDSELADDIEANFTATFVPSIRQRLENDLSGVSLTDTEVTYLM

DMCSFDTISTSTVDTKLSPFCDLFTHEEWINYDYLQSLNKYYGHGAGNPLGPTQGVGYAN

ELIARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFSHDNGIISILFALGLYNGTKPL

SSTTAENITQTDGFSSAWTVPFASRMYVEMMQCQSEQEPLVRVLVNDRVVPLHGCPVDAL

GRCTRDSFVKGLSFARSGGDWGECFA

3 *E. Coli* phytase

E. Coli phytase sequence: accession code B1IZP9
(http://www.uniprot.org/uniprot/B1IZP9)

TLVYCSEGSPEGFNPQLFTSGTTYDASSVPLYNRLVEFKIGTTEVIPGLAEKWEVSEDGK

TYTFHLRKGVKWHDNKEFKPTRELNADDVVFSFDRQKNAQNPYHKVSGGSYEYFEGMGLP

ELISEVKKVDDNTVQFVLTRPEAPFLADLAMDFASILSKEYADAMMKAGTPEKLDLNPIG

TGPFQLQQYQKDSRIRYKAFDGYWGTKPQIDTLVFSITPDASVRYAKLQKNECQVMPYPN

PADIARMKQDKSINLMEMPGLNVGYLSYNVQKKPLDDVKVRQALTYAVNKDAIIKAVYQG

AGVSAKNLIPPTMWGYNDDVQDYTYDPEKAKALLKEAGLEKGFSIDLWAMPVQRPYNPNA

RRMAEMIQADWAKVGVQAKIVTYEWGEYLKRAKDGEHQTVMMGWTGDNGDPDNFFATLFS

CAASEQGSNYSKWCYKPFEDLIQPARATDDHNKRVELYKQAQVVMHDQAPALIIAHSTVF

EPVRKEVKGYVVDPLGKHHFENVSIE

CLUSTAL 2.0.12 multiple sequence alignment

```
tr|Q8J255|Q8J255_ASPFI      ---LAVPASRNQSTCDTVDQGYQCFSETSHLWGQYAPFFSLANKSAISPD 47
A2QIG7|22-467               -----VPASRNQSTCDTVDQGYQCFSETSHLWGQYAPFFSLANKSAISPD 45
B1IZP9|30-535               TLVYCSEGSPEGFNPQLFTSGTTYDASSVPLYNRLVEFK--IGTTEVIPG 48
                                   .* :  . : . .*    :.:  *:.: . *    ..: : *.

tr|Q8J255|Q8J255_ASPFI      VPAGCQVTFAQVLSRHGARYPTDSKGKKYSALIEEIQQN--ATTFEEKYA 95
A2QIG7|22-467               VPAGCHVTFAQVLSRHGARYPTDSKGKKYSALIEEIQQN--ATTFEGKYA 93
B1IZP9|30-535               LAEKWEVSEDGKTYTFHLRKGVKWHDNKEFKPTRELNADDVVFSFDRQKN 98
                            :.   .*:       .  *  .. :.:*     .*:: :  . :*: :

tr|Q8J255|Q8J255_ASPFI      FLKTYNYSLG--ADDLTPFGEQELVNSGVKFYQR--YESLTRNIVPFIRS 141
A2QIG7|22-467               FLKTYNYSLG--ADDLTPFGEQELVNSGVKFYQR--YESLTRNIVPFIRS 139
B1IZP9|30-535               AQNPYHKVSGGSYEYFEGMGLPELISEVKKVDDNTVQFVLTRPEAPFLAD 148
                              :.*:   *   : :  :*  **:..  *. :.     ***  .**: .

tr|Q8J255|Q8J255_ASPFI      SGSSRVIASGNKFIEGFQSTKLKDPRAQP--GQSSPKIDVVISEASTSNN 189
A2QIG7|22-467               SGSSRVIASGNKFIEGFQSTKLKDPRAQP--GQSSPKIDVVISEASTSNN 187
B1IZP9|30-535               LAMDFASILSKEYADAMMKAGTPEKLDLNPIGTGPFQLQQYQKDSRIRYK 198
                             . . .   .::: :.: .:   :       * .. :::   .::    :

tr|Q8J255|Q8J255_ASPFI      TLDPGTCTVFEDSELADDIEANFTATFVPSIR-----QRLENDLSGVTLT 234
A2QIG7|22-467               TLDPGTCTVFEDSELADDIEANFTATFVPSIR-----QRLENDLSGVSLT 232
B1IZP9|30-535               AFDGYWGTKPQIDTLVFSITPDASVRYAKLQKNECQVMPYPNPADIARMK 248
                            ::*    *  : . *. .* .: :. :.   :         *  . . :.

tr|Q8J255|Q8J255_ASPFI      DTEVTYLMDMCSFDTISTSTVDTKLSPFCDLFTHEEWINYDYLQSLNKYY 284
A2QIG7|22-467               DTEVTYLMDMCSFDTISTSTVDTKLSPFCDLFTHEEWINYDYLQSLNKYY 282
B1IZP9|30-535               QDKSINLMEMPGLN-VGYLSYNVQKKPLDDVKVRQALTYAVNKDAIIKAV 297
                            : :   **:* .:: :.  : :.: .*: *: .::        ::: *

tr|Q8J255|Q8J255_ASPFI      GHGAGNPLG---PTQGVGYANELIARLTHSPVHDDTSSNHTLDS------ 325
A2QIG7|22-467               GHGAGNPLG---PTQGVGYANELIARLTHSPVHDDTSSNHTLDS------ 323
B1IZP9|30-535               YQGAGVSAKNLIPPTMWGYNDDVQDYTYDPEKAKALLKEAGLEKGFSIDL 347
                             :*** .     *.   ** :::     ..   .   .:  *:.

tr|Q8J255|Q8J255_ASPFI      ---------NPATFPLNSTLYADFSHD---NGIISILFALGLYNGTKPLS 363
A2QIG7|22-467               ---------NPATFPLNSTLYADFSHD---NGIISILFALGLYNGTKPLS 361
B1IZP9|30-535               WAMPVQRPYNPNARRMAEMIQADWAKVGVQAKIVTYEWGEYLKRAKDGEH 397
                                     ** :  : . : **:::      *::  :.  * ....
```

```
tr|Q8J255|Q8J255_ASPFI      STTAENITQTDG-----FSSAWTVPFASR-------------MYVEMMQC 395
A2QIG7|22-467               STTAENITQTDG-----FSSAWTVPFASR-------------MYVEMMQC 393
B1IZP9|30-535               QTVMMGWTGDNGDPDNFFATLFSCAASEQGSNYSKWCYKPFEDLIQPARA 447
                            .*.  . *  :*     *:: :: . :.:               ::  :.

tr|Q8J255|Q8J255_ASPFI      QSEQEPLVRVLVNDRVVPLHGCPVDALGRCTRDS----FVKGLSFARSGG 441
A2QIG7|22-467               QSEQEPLVRVLVNDRVVPLHGCPVDALGRCTRDS----FVKGLSFARSGG 439
B1IZP9|30-535               TDDHNKRVELYKQAQVVMHDQAPALIIAHSTVFEPVRKEVKGYVVDPLGK 497
                             .:::  *.:  : :**  . .*.  :.:.*  .     ***  .   *

tr|Q8J255|Q8J255_ASPFI      DWAECFA-- 448
A2QIG7|22-467               DWGECFA-- 446
B1IZP9|30-535               HHFENVSIE 506
                            .  * .:
```

```
SeqA Name                     Len(aa)  SeqB Name                     Len(aa)  Score
===================================================================================
1    tr|Q8J255|Q8J255_ASPFI   448      2    A2QIG7|22-467            446      99
1    tr|Q8J255|Q8J255_ASPFI   448      3    B1IZP9|30-535            506      4
2    A2QIG7|22-467            446      3    B1IZP9|30-535            506      4
```

Notes for the sequence alignment for phytases *Aspergillus ficuum, Aspergillus niger* and *E.coli* phytases.

In addition to the phytases above we also used *Trichoderma reesei* phytase but were unable to find the sequence for: this is also called "Finase L", and it is a liquid preparation as purchased consisting of a mixture of that phytase, plus beta-glucuronase and xylanase; this may account for the poor aqueous activity in water (we did not test it in organic as it was a liquid preparation) as the phytase units may be low compared to the other phytase preparations tested in our work.

For the first 2 phytases, we found 99% homology, and this may account for the very similar aqueous activity of the 2 enzymes, the small difference in the HMDS condensation being primarily the amount of enzyme and/or enzymatic activity in the liquid preparation.

Lastly, E. Coli phytase gave a very surprisingly low homology sequence compared to the other 2, only 4%; This large difference might explain again the different behaviour of E Coli respect to the other 2, i.e. good activity only in organic media.

In the alignment, * means identical residues, : means highly similar residues, and . means residues with similar functionalities. Clustal W was used to align the 3 sequences (the order is always 1 *Aspergillus ficuum* phytase; 2 *Aspergillus niger* phytase; 3 *E. coli* phytase)