

A CHO-binding domain I

```
Cdd:pfam 7  GTFEDGGLGGWKAR-GGSVKAT-VDVTSN-GTYSLKVSGR-----TA---TAIIDITIRLEKGTTY 60
              * * ^ **^* * ^ ^ **          ^* * *
PA2783 302  GDGEFASMEGWRAWSGNAQLSL-VNVAKGC-RDNALLVDVR-----GF---DLLVRPIAPLRAGSGY 358
Pmendo 299  _NGEFTTSDWRALFGQSDLSL-VNVALNC-RDNALQMDAR-----GF---DILATPVSGLRDGVQY 354
Hcheju 303  _DGDFFNELSSWTSYSNAGDLGQ-AYFKKDCGRDYILEITNR-----SAYYAGAYQKLTGKLGAGVEY 362
Pmucil 38   HDFEDA AVAGWTRGGVEILAA-SPGAAHS-GVYGLQVDGRTKGWHGP---SLEVTPL--MRAGQSY 97

Cdd:        61  TVSEFWKASSGPQ--TISVTLQITDASGNYDTVADEKVVLT---GEWTKLEGTFTIPGTASTVELYV 127
              ^* * ^ *^ ^ * * *          ^* * ^ *^^** * * *
PA2783 359  RLSGKVMKKAANTRETVRMALLSERADGALAYNPAQSVELSVSGNEFSRLEKTFDYRPAADQRNLYV 425
Pmendo 355  RLRANSMLKARN SRENVR LAIQEGLDGRITYSSEQSASLSVSGNEFSRLEKTFYRTPANARNLYV 422
Hcheju 363  KLTA KLGLGENG SRENMRV-VFKVSDDSGTHYQYLS--RASFTNSEMSKYERTFTLDAKGTIRDVGM 426
Pmucil 98   VVSGWLKLPAGSPNTKVYLSLQHS LAAGE-QYEQIASAAVTS SG--VWKIEAQYKLR EAANKLSVYF 161

Cdd:        128  ELG-PDST 134
              ^ **
PA2783 426  AVW-SDSG 432
Pmendo 423  AVW-SDSG 429
Hcheju 427  LIYGASAG 479
Pmucil 162  EAP-DQPA 186
```

B CHO-binding domain II

```
Cdd:pfam 7  GTFEDgGLGGW----KARGGSVKA--TvDVTs-----HnGTYSLKVSGRTATWDGAIIDITIRLEKGTTYTVSFWV 71
              ** *^***          * ^* * * *          * *^ * * * * * * * * *
PA2783 461  YDFES-GIGGW----SGVHASTRA--T-RVAS-----A-GRLEAYQRRYAGTGASTSLLGNLEAGRTYAFSADV 523
Pmendo 455  WNFES-GIGGW----SGYHGSARV--S-TYAS-----S-GRQALEAYQREYEGSGASVSLGNLEAGKRYRIEADL 517
Hcheju 472  EGFEN-GTRGW----SN-YFGTRLYLS-DTAA-----S-GKYSLLSSNR YRWYSGPGLNAHGLEAGKPYQVSADV 534
Clen-1 39   YNFDG-GTEGW IARGDAQVTTSS--S-LTID-----S-TNHCLYVTNRTSNWHGVSKELKDTLKAGETYTFSVYV 104
Clen-2 197  WDFES-GLGDW----S-VRSGGEA--T-AEVSTTAPHG-GNQCAFVSGRTQDWNGIQVNL TNSAKKGGKYTFEAWV 262

Cdd:        72  -----KAS--SGPQ TISVTLQITDASGNYDTVADEK-VVLTGEWTKLEGTFTI-PGT-ASTVELYVELG 130
              ^ *          ^ *^ * *          * * * * * * * * * * ^ *
PA2783 524  RVGDG-RGS--QAMTYAYLYLESQGRPEY-LPLGYK-VVENGRWASLRGQVQL-PKGPIKRAELMILSG 586
Pmendo 518  SIGRS-TSV--SAVAYAYLYVEDSNGRQY-LPLGQR-NTRGGSSK LQHEVQL-PAGPLRAELLLL-- 578
Hcheju 535  YLASS-RRS SDAELWVY-YVDNSG--GHW-KKLGQ-EIATSQWQNI SGEINLSPNGTISQLRLHVIGA 598
Clen-1 105  KYDEG-ADT--ESLS---LTLETGLEENGY-ITIASNSAVSKGEWTQLTGEVTV-PEAATK-ANVYMEAA 165
Clen-2 263  KYDAADEGT--EQFV LALQYDAESESSTQY-KWITNA-DVKNGEWTKLQGEYTL-PK----- 314
```

Additional file 3 - The predicted PA2783 protein carries two carbohydrate-binding modules

Interrogation of the non-redundant databases at NCBI (<http://www.ncbi.nlm.nih.gov/>; accessed 06/19/2013) using BLASTP revealed homology with the CBM_4_9 family (Cdd:pfam02018) of diverse CHO-binding proteins. CHO-binding domain I (A) and domain II (B), have different aa sequences but both were strongly homologous to the two CHO-binding modules of the *Pseudomonas mendocina* (Pmendo) carbohydrate-binding CenC domain-containing protein and the Ni,Fe-hydrogenase I small subunit of *Hahella chejuensis* (Hcheju). For the pfam, identical aa are indicated by * and similar aa by ^; for bacterial proteins, identical aa are shown in red, similar aa in blue, and non-similar aa in black; Pmucil, *Paenibacillus mucilaginosus*; Clen-1 and Clen-2, *Cellulosilyticum lentocellum* CHO-binding domains I and II. Percentages of aa identity and similarity are shown in Additional file 4.