

Supplementary methods

Among three OCD-like genes from *T. litoralis* DSM 5473, *T. sibiricus* MM 739, and *K. cryptofilum*, clustered with putative T3LHyp dehydratase genes, only the gene from *T. litoralis* DSM 5473 consisted of two separate open-reading frames (ORFs): OCC_00362 and OCC_00367 (*red* and *blue* in the following scheme). On the other hand, all known OCD/CRYM superfamily enzymes are encoded by a single ORF, such as genes from *T. sibiricus* (TSIB_0634) and *K. cryptofilum* (Kcr_0796) (see Fig. 1D). In fact, we found that when a single “A” is inserted between A227646 and T227647 in the original genome sequence of *T. litoralis* DSM 5473 (underlined lower case letter), the separate ORFs fuse into a single ORF, and the putative amino acid sequence of (*green*) is clearly homologous to that of the TSIB_0634 gene (*orange*). There are two possibilities from these preliminary analyses. An error in the genome sequence was not found in the sequencing. Although there is another possibility that this OCD-like protein from *T. litoralis* DSM 5473 forms a heterooligomeric structure, we synthesized the *TLLhpI* gene as a single ORF by sequential steps of PCR [1] using sense and antisense primers (Table S1) and genome DNA of *T. litoralis* DSM 5473 as a template.

1. Penning TM & Jez JM (2001) Enzyme redesign. *Chem Rev* **101**, 3027-3046.

OCC_00362

OCC_00367

T A R E P V V K G D G L G K A R T S T P *

(M) R K S K L V

ACGGCAAGAGAACCAGTAGTTAAGGGGGATGGATTAGGGAAGGCACGCACATCAACTCCGTAGGCTGGGTGGGAGAGATGCCAGGGAACCGATTCCGAAACAGTGAGAAAGTCTAAGCTGGTT

ACGGCAAGAGAACCAGTAGTTAAGGGGA_aTGGATTAGGGAAGGCACGCACATCAACTCCGTAGGCTGGGTGGGAGAGATGCCAGGGAACCGATTCCGAAACAGTGAGAAAGTCTAAGCTGGTT

T A R E P V V K G E W I R E G T H I N S V G W V G R D A R E L D S E T V R K S K L V

ACTGCAAAGAACCCTGATTAAGGGAGAATGGATCAAAGAAGGCACGCACATAAATTCTATAGGGTGGATGGGAAAAGATGCCAGAGAACCGATTCCGAAACAGTAAGAAAGTCCAAACTGGTC

T A K E P V I K G E W I K E G T H I N S I G W M G K D A R E L D S E T V R K S K L V

TSIB_0634