Biochemical differences and similarities between the DEAD-box helicase orthologs DDX3X and Ded1p

Deepak Sharma, Andrea A. Putnam & Eckhard Jankowsky

SUPPLEMENTARY DATA

```
CLUSTAL O(1.2.4) multiple sequence alignment

SP|P06634|DED1_YEAST    --------MAELSEQVQNLSI---------NDNNENGYVPPHLRGKPRSARNNSSNYNNN 43
SP|O00571|DDX3X_HUMAN   MSHVAVENALGLDQQFAGLDLNSSDNQSGGSTASKGRYIPPHLRNREATKG----FYDKD 56
                               *.:*. .*.:           .  .:. *:*****.:  :        *:::

SP|P06634|DED1_YEAST    NG--------------GYNGGRGGGSFFSNNRRGGYG-----------NGGFFGGNNGG 77
SP|O00571|DDX3X_HUMAN   SSGWSSSKDKDAYSSFGSRSDSRGKSSFFSDRGSGSRGRFDDRGRSDYDGIGSRGDRSGF 116
                        ..              . ...** .****:.  *. *           . *  *...*

SP|P06634|DED1_YEAST    SRSNGRSGGRWIDGKH-----VPAPRNEKAEIAIFGVPEDPNFQSSGINFDNYDDIPVDA 132
SP|O00571|DDX3X_HUMAN   GKFERGGNSRWCDKSDEDDWSKPLPPSERLEQELFS------GGNTGINFEKYDDIPVEA 170
                        .: :  ...** * ..       * * .*: *  :*.           .:****::******:*

SP|P06634|DED1_YEAST    SGKDVPEPITEFTSPPLDGLLLENIKLARFTKPTPVQKYSVPIVANGRDLMACAQTGSGK 192
SP|O00571|DDX3X_HUMAN   TGNNCPPHIESFSDVEMGEIIMGNIELTRYTRPTPVQKHAIPIIKEKRDLMACAQTGSGK 230
                        :*:: *   * .*:.  :. :: **:*:*:*:*****::**: : *************

SP|P06634|DED1_YEAST    TGGFLFPVLSESFKTGPSPQP---ESQGSFYQRKAYPTAVIMAPTRELATQIFDEAKKFT 249
SP|O00571|DDX3X_HUMAN   TAAFLLPILSQIYSDGPGEALRAMKENGRYGRRKQYPISLVLAPTRELAVQIYEEARKFS 290
                        *..**:*:**: :. **.       :.:* : :** ** :::::*******.**::**:**:

SP|P06634|DED1_YEAST    YRSWVKACVVYGGSPIGNQLREIERGCDLLVATPGRLNDLLERGKISLANVKYLVLDEAD 309
SP|O00571|DDX3X_HUMAN   YRSRVRPCVVYGGADIGQQIRDLERGCHLLVATPGRLVDMMERGKIGLDFCKYLVLDEAD 350
                        *** *: ******: **:*:*::**** ********* *::**** .*   *********

SP|P06634|DED1_YEAST    RMLDMGFEPQIRHIVEDCDMTPVGERQTLMFSATFPADIQHLARDFLSDYIFLSVGRVGS 369
SP|O00571|DDX3X_HUMAN   RMLDMGFEPQIRRIVEQDTMPPKGVRHTMMFSATFPKEIQMLARDFLDEYIFLAVGRVGS 410
                        ************:***:   * * * *:*:******* :** ******.:****:******

SP|P06634|DED1_YEAST    TSENITQKVLYVENQDKKSALLDLLSAS-TDGLTLIFVETKRMADQLTDFLIMQNFRATA 428
SP|O00571|DDX3X_HUMAN   TSENITQKVVWVEESDKRSFLLDLLNATGKDSLTLVFVETKKGADSLEDFLYHEGYACTS 470
                        *********::**:.**:* ***** *:  .*.***:****: **.* ***  ::. .*:

SP|P06634|DED1_YEAST    IHGDRTQSERERALAAFRSGAATLLVATAVAARGLDIPNVTHVINYDLPSDVDDYVHRIG 488
SP|O00571|DDX3X_HUMAN   IHGDRSQRDREEALHQFRSGKSPILVATAVAARGLDISNVKHVINFDLPSDIEEYVHRIG 530
                        *****:* :**.** **** : :*************** **.****:*****::*******

SP|P06634|DED1_YEAST    RTGRAGNTGLATAFFNSENSNIVKGLHEILTEANQEVPSFLKDAMMSAPG-------SRS 541
SP|O00571|DDX3X_HUMAN   RTGRVGNLGLATSFFNERNINITKDLLDLLVEAKQEVPSWLENMAYEHHYKGSSRGRSKS 590
                        ****.** ****:***..* **.*.* ::*.**:*:*****:*::    .        *:*

SP|P06634|DED1_YEAST    NSRRGGFGRNNNRDYRKAGGASAGGWGSSRSRDNS----------FRGGSGWGSDSKSSG 591
SP|O00571|DDX3X_HUMAN   SRFSGGFG---ARDYRQSSGASSSSFSSSRASSSRSGGGGHGSSRGFGGGGYGGFYNSDG 647
                        .  ****   ****::.***:..:.***: ..        **.*:*. :*.*

SP|P06634|DED1_YEAST    WGNSGGSNNSSWW-- 604
SP|O00571|DDX3X_HUMAN   YGGNYNSQGVDWWGN 662
                        :*.. .*:. .**
```

Supplementary Figure S1. Sequence alignment of human DDX3X and yeast Ded1p.
UniProt IDs for DDX3X and Ded1p (DED1) are shown on the left. Conserved residue positions
are indicated by an asterisk. Colons and periods indicate similar amino acid residues.
Unstructured regions are highlighted in Red. Unstructured regions were predicted by a
Supervised Machine Learning technique implemented in PrDOS using the default parameters
(*ref.53*). Amino acid sequences for DDX3X and Ded1p were converted to a position-specific
score matrix (PSSM), which was then used to predict unstructured regions according to: (i)

calculating the probability of absence of structure through comparison of DDX3X/Ded1p sequences with crystal structure information about the non-redundant protein dataset from the Protein Data Bank (PDB) and (ii) alignment of DDX3X/Ded1p sequences to relevant structures determined in the previous step ($e < 10^{-3}$) (False positive rate was set to 1%). The output was a disorder profile based on the unstructured region propensity with respect to template crystal structure of DDX3X (PDB ID: 2I4I). The identical number of amino acids in unstructured regions for both proteins is a result of the combinatorial template based prediction algorithm that uses crystal structures of DDX3X fragments only, given that no crystal structure for Ded1p is available.

Supplementary Reference

53   IshidaT et al, (2007) PrDOS: prediction of disordered protein regions from amino acid sequence; *Nucleic Acids Res*. 35 (Web Server issue): W460–W464.

DDX3X						
K_m (µM)	1019 ± 52	1517 ± 71	860 ± 38	970 ± 41	1543 ± 65	1243 ± 52
k_{cat} (min^{-1})	78 ± 3	36 ± 1	63 ± 2	75 ± 2	2.5 ± 0.1	2.0 ± 0.1
k_{cat}/K_m (mM^{-1}min^{-1})	0.077 ± 0.011	0.023 ± 0.009	0.074 ± 0.014	0.077 ± 0.017	0.002 ± 0.001	0.001 ± 0.001
$K_{1/2\ [RNA]}$ (µM)	2.3 ± 0.3	3.4 ± 0.6	3.4 ± 0.6	2.1 ± 0.2	2.8 ± 0.5	3.3 ± 0.6
Ded1p						
K_m (µM)	850 ± 41	1100 ± 56	760 ± 34	1000 ± 47	1743 ± 56	1349 ± 38
k_{cat} (min^{-1})	75 ± 2	30 ± 2	61 ± 2	29 ± 2	2.0 ± 0.1	2.5 ± 0.1
k_{cat}/K_m (mM^{-1}min^{-1})	0.088 ± 0.020	0.027 ± 0.008	0.080 ± 0.010	0.029 ± 0.013	0.001 ± 0.0002	0.001 ± 0.003
$K_{1/2\ [RNA]}$ (µM)	2.7 ± 0.4	2.1 ± 0.4	1.7 ± 0.5	1.6 ± 0.4	2.9 ± 0.6	1.6 ± 0.3

Supplementary Table S1. Michaelis-Menten parameters for ATPase activity of DDX3X and Ded1p with different substrates. Rates of ATP hydrolysis by DDX3X and Ded1p were measured as a function of ATP concentration in the presence of different RNA substrates, shown in the top row. Grey lines represent RNA. Black lines indicate DNA. Each duplex substrate is 16 bp with a 25 nt tail, 3' to the duplex region. Single stranded substrates are 41 nt. Values for k_{cat} and K_m were obtained by fitting of ATP titrations (9 individual ATP concentrations up to 3 mM ATP per substrate, 450 nM protein, 4.5 µM substrate, except the ssDNA and DNA duplex, which were 6 µM) to the Michaelis-Menten equation. Reported errors represent the 95% confidence interval from these data fits. Values for $K_{1/2}$ [RNA] were obtained from measurements of rates of ATP hydrolysis at 3 mM ATP, 450 nM protein and titrations of the respective substrates up to 4.5 µM, except the ssDNA and DNA duplex, which were 6 µM, and subsequent fits to the Michaelis-Menten equation. Values of $K_{1/2}$ [RNA] correspond to the calculated K_m values of the data fits, reported errors represent the 95% confidence interval from these data fits.