

May 5, 2014

PLOS ONE

Title: Structural insights of the ssDNA binding site in the bifunctional endonuclease AtBFN2 from *Arabidopsis thaliana*

Supplementary materials

Tsung-Fu Yu ^{a,b,†}, Manuel Maestre-Reyna ^{b,†}, Chia-Yun Ko ^c, Tzu-Ping Ko ^b, Yuh-Ju Sun ^a, Tsai-Yun Lin ^a, Jei-Fu Shaw ^{c,d,*}, and Andrew H.-J. Wang ^{a,b,c*}

^aInstitute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu 30013, Taiwan

^bInstitute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan

^cAgricultural Biotechnology Research Center, Academia Sinica, Taipei 11529, Taiwan

^dDepartment of Biological Science and Technology, I-Shou University, Kaohsiung 84001, Taiwan

†TFY and MMR contributed equally to this work.

*To whom correspondence should be addressed:

JFS: shawjf@isu.edu.tw

AHJW: ahjwang@gate.sinica.edu.tw

```

AtBFN2(28-290) WGKEGHEIICKIAQTRLDETAAKAVKELLPESEAEGDLSLCLWADRVK--FRYHWSPLH 58
BEN1(24-288) WGKEGHYMTCKIADGFLTSEALTVGKALLPSWANGELAEVCSWADSSQR--FRYRWSRSLH 58
ABN1(29-305) WSKEGHIILTCRI AQNLLLEAGPAHVVENLLPDYVKGDLSALCVWPDQIRHWYKYRWTSSHL 60
AtBFN1(29-305) WSKEGHIILTCRI AQNLLLEAGPAHVVENLLPDYVKGDLSALCVWPDQIRHWYKYRWTSSHL 60
ZEN1(26-303) WSKEGHVMTCCI AQELLSPDAAHVAQMLLPDYVKGNLSALCVWPDQIRHWYKYRWTSSHL 60
CEL1(23-296) WSKEGHVMTCCI AQDLEPEAAHAVKMLLPDYANGNLSL CVWPDQIRHWYKYRWTSSHL 60
HBN1(25-300) WSKEGHTMTCQI AQALLDSEAAEAVRNLLPDYVNGDLSALCVWPDQVRHWYKYRWTSSHL 60
TBN1(26-302) WSKEGHVMTCRI AQGLLNDEAAHAVKMLLPDYVNGDLSALCVWPDQVRHWYKYRWTSSHL 60
P1 nuclease(1-270) WGALGHATVAYVAQHYVSPAASWAQGLGSSSSSYLASIASWADEYRLTSAKWSASLH 60
BcPLC(39-283) WSAEDKHKEGVNSHLWIVNRAIDIMSRNTTLVKQDRVAQLNEWRTLENGIYAADYENPY 60
Consensus WSKEGHVMTCI*IAQ*LLEPEAAHAVK*LLPDYVNGDLSALCVWPDQIRHWYKYRWTSS* LH
AtBFN2(28-290) YINTPD----ACSYQYNRDCKDESGEKGRCVAGAIYNYTTQLLSYKTAASSQSQYNL TEA 114
BEN1(24-288) FADTPG----DCKFSYARDCHDTKGNKNVCVVGAINNYTAALQD----SSSP--FNPTES 108
ABN1(29-305) YIDTPDE----ACSYEYSRDCHDQHGL EDMCV DGAIQNF TSQLQHYGEGTSDR-RHNMTEA 116
AtBFN1(29-305) YIDTPDQ----ACSYEYSRDCHDQHGLKDMCV DGAIQNF TSQLQHYGEGTSDR-RYNMTEA 116
ZEN1(26-303) FIDTPDD----ACSFYTRDCHDSNGMVDMCVAGAIKNFTSQLSHYQHGTSDR-RYNMTEA 116
CEL1(23-296) FIDTPDQ----ACSFYQRDCHDPHGKDMCVAGAIQNF TSQLGHFRHGTSDR-RYNMTEA 116
HBN1(25-300) FIDTPDQ----ACSFYSRDCHDQHGLENM CVAGAIQNF TSQLSHYREGTSDR-RYNMTEA 116
TBN1(26-302) FIDTPDK----ACNFDYERDCHDQHGVKDMCVAGAIQNF TQLSHYREGTSDR-RYNMTEA 116
P1 nuclease(1-270) FIDAEDNPPTNCNV DYERDCGSSG----CSISAIANYTQRVSDSSLSSENH----AEA 110
BcPLC(39-283) YDNSTFA-----SHFYDPDNGKTYIPFAKQAKETGAKYFKLAGESYKKNKDMK----QAFFY 112
Consensus FIDTPDQ----ACSFYSRDCHDQHGLKDMCVAGAIQNF TSQLSHYREGTSDR-RYNMTEA
AtBFN2(28-290) LFLVSHFMGDIHQPLHVSYASDKGGNTIEVHWYTRKANLHHIWDSDNI IETAEADLYNSAL 174
BEN1(24-288) LMF LAHFVGDVHQPMHCGHVDDLGNGT I KLRWYRRKSNLHHVWSDVITQTMKDFDQDKDQ 168
ABN1(29-305) LFLFLSHFMGDIHQPMHVGF TSEGGNTIDL RWRH KSNLHHVWDR E I I L T A L K E Y D K D L 176
AtBFN1(29-305) LFLFLSHFMGDIHQPMHVGF TSEGGNTIDL RWRH KSNLHHVWDR E I I L T A L K E Y D K D L 176
ZEN1(26-303) LFLVSHFMGDIHQPMHVGF TDEGGNTIDL RWRH KSNLHHVWDR E I I L T A A S E L Y D K D M 176
CEL1(23-296) LFLFLSHFMGDIHQPMHVGF TSDMGGNSIDL RWRH KSNLHHVWDR E I I L T A A A D Y H G K D M 176
HBN1(25-300) LFLFLSHFMGDIHQPLHVGF TDEGGNTINL RWRH KSNLHHVWDR E I I L T A L K D Y Y E K N T 176
TBN1(26-302) LFLFLSHFMGDIHQPMHVGF TSDAGNSIDL RWRH KSNLHHVWDR E I I L T A A K D Y Y A K D I 176
P1 nuclease(1-270) LRFLVHF I G D M T Q P L H D - E A Y A V G G N K I N V T F D G Y H D N L H S D W D T Y M P Q K L I G G H A L S D A 169
BcPLC(39-283) LGLSLHYLGDV NQPMHAANFTNL S----YPQG FHSKYENFVDITKDN YKVT D G N G Y W N W K G 169
Consensus LFLFLSHFMGDIHQPMHVGF TSEGGNTIDL RWRH KSNLHHVWDR E I I L T A * K D Y Y D K D L
AtBFN2(28-290) EGMVDALKKNITTEWADQVKRWETC---TKKTACPD IYASEG IQAACDWAYKG---VTE 227
BEN1(24-288) DAMIESIQRNITD-DWSSEEKQWETCR--SKTTTCAEKYQESAVLACD-AYEG---VEQ 221
ABN1(29-305) DLLQEDLEKNI TNGFWDDDLSSWTEC---IDL IACPHKYASES IKLACKYGYEG---VKS 230
AtBFN1(29-305) DLLQEDLEKNI TNGLWDDDLSSWTEC---NDL IACPHKYASES IKLACKWGYEG---VKS 230
ZEN1(26-303) ESLQKA IQANFTHGLWSDDVNSWKDC---DDI SNCV NKYAKES IALACKWGYEG---VEA 230
CEL1(23-296) HSLLQDIQRNFTEGSWLQDVE SWKEC---DDI STCANKYAKES IKLACN WGYKD---VES 230
HBN1(25-300) DLLLQDIKGN YTDGVWSDDVSW EHC---DDLASC INKFAVES INIACKWGYKD---VEP 230
TBN1(26-302) NLLEEDI EGNFTDGIWSDDLASWREC---GNVFCVNKFATES INIACKWGYKG---VEA 230
P1 nuclease(1-270) ESWAKTLVQNI ESGNYTAQAI GWI KGDNI SEPI T T A T R W A S D A N A L V C T V V M P H G A A A L Q 229
BcPLC(39-283) -----TNPEW I H G---AAVVAKQDYSGI VNDNTKDFVYKA---AV 204
Consensus DLLQEDI*KNITDGLWSDDVSWEEC---DDLI*C*NKYASES IKLACKWGYKG---VES
AtBFN2(28-290) GDTLEDEYFY S R L P I V Y Q R L A Q G G V R L A A T L N R I F G-----263
BEN1(24-288) DDTLGDEYFYKALPVVQKRLAQQGLRLAAI LNRIFSGNGRLQSI-----265
ABN1(29-305) GETLSEDYFNTRM P I V M K R I V Q G G V R L A M I L N R I F S D T H--AGVAATCG 277
AtBFN1(29-305) GETLSEYFNTRLP I V M K R I V Q G G V R L A M I L N R V F S D D H A I A G V A A T--277
ZEN1(26-303) GETLSDDYFDSRMP I V M K R I A Q G G V R L S M I L N R V F G S S S S L E D A L V P T--278
CEL1(23-296) GETLSDKYFNTRM P I V M K R I A Q G G I R L S M I L N R V L G--SSADHSLA--274
HBN1(25-300) GVTLADDYFDSRMP I L M K R I A Q G G V R L A M I L N R V L G-ESD-EGLATPT-276
TBN1(26-302) GETLSDDYFNSRLP I V M K R V A Q G G I R L A M L L N N V F G-ASQQEDSVVAT-277
P1 nuclease(1-270) TGDLYPTYYSV I D T I E L Q I A K G G Y R L A N W I N E I H G S E I A K-----270
BcPLC(39-283) SQEYADKWRAE V T P M T G K R L M D A Q R V T A G Y I Q L W F D T Y G D R-----245
Consensus GETLSDDYFNSR*PIVMKRI A Q G G V R L A M I L N R V F G-ES*EG-----

```

Figure S1. Alignment of S1/P1-type nucleases. Amino acid sequences of AtBFN1

(UnitProt: Q9SXA6), AtBFN2 (Q9C9G4) from *Arabidopsis thaliana*, BEN1

(O81958) from *Hordeum vulgare*, ABN1 (E3PQH5) from *Fourraea alpina*, ZEN1

(O80326) from *Zinnia elegans*, CEL1 (Q9LL59) from *Apium graveolens*, HBN1

(B4ERM5) from *Humulus lupulus*, TBN1 (Q0KFV0) from *Solanum lycopersicum*,

P1 nuclease (P24289) from *Penicillium citrinum* and BcPLC (P09598) from *Bacillus cereus* were aligned using the ClustalW program. Binding pocket 1 (yellow), pocket2 (green), pocket3 (cyan) and pocket4 (pink) are labeled. The tri-metallic zinc active site is labeled in red and (*). Finally, the highly variable Tyr-site amino-acids are labeled with black outlines.

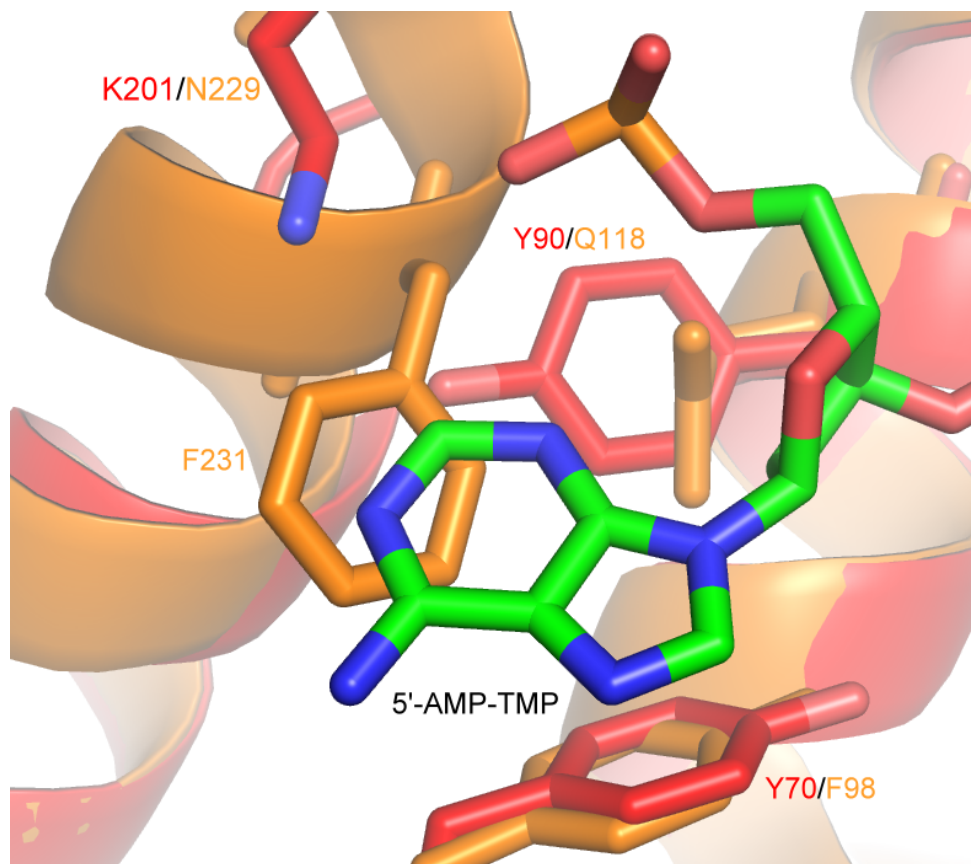


Figure S2. Comparison of secondary binding site between AtBFN2/A5T (red) and TBN1 (3SNG, orange) with conserved protein residues. Tyr90 in AtBFN2 is replaced by a glutamine in TBN1. However, an extension of α helix 11 in TBN1, which AtBFN2 is missing, supplies Phe231.

Table S1. Primers used for PCR-amplification of constructs applied in this study.

For the overexpression construct of

ENDO2

Sequences of primer pairs (5' → 3')	Enzyme site
--	--------------------

B2-11F:	BamHI
----------------	--------------

CGGGATCCATGGCAAACCAA

B2-12R:	SacI
----------------	-------------

CGAGCTCTCAGTGGTGGTGGTGG

TGGTGACCGAAAATCCT
