

## Appendix 4a

Complete protein sequences for Amphioxus 2A-like (NCBI accession number is provided in bold)

### XP\_002204853.1

MGIDLSTYRARIGCFNSNPGHDSLHCWAEASCLLVWMVISQLMLKLAG**DVEENPGPQERYLGKTMGTQPMRDEMKR**  
EEWKILGKAQPLWTLDLSNQNHKDLPDEVFELEELAALKLTYNESISLSNKLVKLTKLNLKVLCLENCNLDKLPPVV  
LKLSHLQVLDISKNKAISLPKMLKKLKLKVLKLRDCDLVTIGRQIFQQESQLEELDSGNMQIDLPDELRTLK  
NIRVLRLNRAGMTTVPPAVLELSQLEKLDLSGNKQIKLSDQLLGLTNLKVLRLSRTEMASVPEVVWLTHLEELH  
LLSNPLQTLCSVKGQLSRIKRLDLSNCHLRTLPPEVGTLTQLERLKVANNRALQTLPGELWQVTNIKRLDLSNCQ  
LHTLPPEVGTLTQLEWLDSLNTLQTLPRELGHVTNIKRLDLSHQLHTLPPQVGKLTBLKWLKVKNPLQTLPG  
ELGQVASIKHLDLSNCWLHTLPPEVGTLTQLERLKVANNPLQTLPGELWQVTNIKRLDLSSCWLDTLPPEVGTLT  
QLEWLSLQGNPLQMLPKQIGQLTAIKHNLNSFCQLHTLPPEMGLKQLEWLSLQGNPLQMLPKQVENLTHIKWMN  
LSHCRLQMLPPEFGKLTQLERLYLSCNGEQLTLPTRQLTNIKHLDLSNCISLQTLPPPEVGELKHVEYLRLSSNPLQ  
KLPPEVRHTNIKHLDMSNCRLNELPIEVGTMQLRQLDLRYNQLQMLPVEITQHINLYHLDVRGNPLIRPPAEV  
CSQGMVAVRQYFEELESGEAVSTHLKVVVLGKELAGKTSLIQTLLSGDSSLTQQEDRTHGVEITQWKPDNNITFE  
VYDFGGHDAYHLTHQFLTQDSLHLLTVDLQTYCCTEQDYKNAVGFWLDTLNARVPGAMTIVCSKSEGSVDIEK  
KTHDIEVKFKEQQDTWKRNLERQLIITERAENQEDLKQQFEDIQQLMTRPLQFKGLCCVSSAGPTGETLTAHVL  
ELAKNTQLFQTHKGILPQTWDFEQQLRNLRDNEAKWLMPADYQQQAQLAGLPTDKLEPVLSYLLQVGTILRYTD  
IPELRFVFHDPPALIEIFKGLFHNTSDLFSTLDIYFTPLDLDKFSDLDDRGLLKKVMKHLPPDTPSNIVT  
SLMQHYDLCFEIDECKTKASDEETRQYMIPWYLHKTMQEKWQHNWPQHVPEEQLQLMCNIKGFCPRGLFPRFSVG  
IHSLIKDRVDWKDSVMAYRQDFPVLALCSKPGTDYTITIATRGRILAQADEMWGVVYPLLEVLVQLLQKWPGVLY  
SLHVTCAHCIKARLDNPHQYNLRDQTADDRIVRPCKADFRATVSTDLYRPLQQNKFQRYKMKTDPRGLALII  
SNINFMDPETDRQGGNNDLENIWKICDKLKKAENKDLTGQEIAADTLQEVSERDHSCYDCFVLFLLSHGDEKGV  
LGTDEDRIISVDSITSIQLCKSLVGKPKLFFIQACRGDNFDEGVEYMTDDIKSPSRVGAELSESIMATDSAS  
HQSTTPTASDGNNAKRKTIPTASDRLVMNATVQGQVAWRDKNTGSPFIQALADVITEHADTVTLQDMATLVNEKV  
ADYEADADDDYKQQSEVSHSLRKKLYFFPETAVD

### XP\_002218646 (not tested)

MSSLVNCRKRIA FMTHRLLKCLFLLGIVAVTVGAEIPPVYQTAEGWSCHKCPRGQFVASDCASDDPASTDCRNC  
TGDRFQPNEWTSEKRCYNCDECDPGLGDSLGGDPHTPCTIFENAVCRCPVAFFWSQAQRCEEVTECPGPQGVTH  
PATIKTDTVCEPCREEAYSAKSSAEKPCAACLICEEWQREASPCNQTHNRCVNVSMATGIPPTPKHPQEVSDE  
DSQTDNTVAIVAPVAVTVDFPSNPILSFWSFEAKMCYCSYVILVLLLLMAG**DVEVNPGP**GQDPFSTSNNFKKI  
AGLQHEVQQAQQQQQQQQQQQQQQQQMYQQVLYIPVAQQTQHVPQAALNEERQRRLAEESARNLERQLRECRRQVQL  
AERALREERGVAARANERVAEVERRVVAELEQRVKSKKEVRQKRSSFKKMGSTKPSSQLTASGNKKRKQAYKARLN  
HTFRNFHTDVKKARVELFLEGHNKPHIWKWEEREPSQPQPSQDSDKTREVQKWLVRDKFTISDEGYHEIRMMDQ  
ANVPPPLYRVIIEERVRQNKTDIDIITEANNMARRPVRELLKELLKHPKYGGTVSDQISVRFSGDGRQVTRNNRIGA  
VMGTLRIVPNRNTINEDTPSTRLHTINEEASIFVYEGGEDHEVQKQTAATVYQEIKDIRENGLIIGSKTVKVQW  
YLTADWKFLAILLGLNQANSNSFCLWCHCTKDQICNFDILNWDIQRKVGHDHTCLIGQREHKGHILTPVDFKFDA  
IIPDVLHMGMRIRGKLFNQVVVWAINQRRTKELAAAMKETGVPFRLVEGTGDDGKGSIKTWTQLNQKLERVFHK  
LNLHAVLDRWSANGISVASLNVAKLTELQKHGEIPSKSKKADLVRLLTDALSDQDFEELAVAMKAAPGAAG  
HLTHTVFKQKARDWGTFFRRVTYDEHVTPIHAMVYHVPQFLEYHFINDLSCESVEQVMYEYERDLYAVVHELD  
RKKANMGPNAAQRATMQEEQQGEEEVFPQGEDEDEEGDEEGDEEGDEEKGRRKNLYRRVRRKMRMRKSS  
KKMMMMTHNVMLYTNCQSF

### **XP\_002224141**

MGVDLATYRARIGCFQNPTPRSVRCWSQASALLVGVILAQIWDKVTGDGQGEVVSSPSDWVLPAHQTEVQTYQL  
HCHPQSRCFYVMMSCLLWFMMVHKLLLQAG**DIEPNPGETLRLQSLQTLDVSVNNTIELPGKLGDLSLTVLNVG**  
SNHLKDLPESTKLVNLTSVDISNDLKAIPEEMFEVNMTEFNIAGNKIETLPNTICELQNLEYLSLEENNIRV  
LPTQFGKLHKLHHFGKEQIQDNPMEQPPIEVFEQGMEGVVAYFEELNLSKAVEVPRVKALLGEVAAGKTSLCNA  
IRLGKSNLTDIADRTAGIEVHPTRLQDSIQVLMHDFGGHKIYHLTHQFFFSNAALYLIIVVDLKAFKRSSFAAAVK  
YWLNVIKARIIDSKPVLRIVGTHIDLCDSEEIKEADLIMKKMKQIENIEITLRKHLSIDHALSRAGTPPSDWC  
YALEGFDTQSLRKRRKFIEHLIDNRPIFPEHIDLRRKRALEECGKSPAEEVMMEQPRYSEQEPDISVRNNNTV  
HATCVPSCDCNAQNYDDTENAWERFRSMALDMQPQGEGNRLATHLMEGLKLRPDVKDGNCLFYAVFDQLVRTGS  
TAISHSQLRQDVVGYLKGKPHNEHGDLRAFPNEDWEDYLQQMSRDGVWDHIVLQAMASMLGRDIRIVSSIDA  
ENYTTILSPMGNQQVTTGPPLLLGHYAENHYASLDVCQGSAVLLISDEYGTSGGGISTINHOLAQLLRCFGAVVY  
CTVLTWTTEDEEDARKDGVHLIRPSVQAGDARNPSLDWLTYDHRIRYPNLPKDVGYIVVHGAVTDKAENIKNSR  
YPAQAKLYTFNHTIPEDTEHYKGGRRAIKAKEKETDLLKIADVADAVFSGRRIYNHFETMYKGDKPKAHHLFLP  
KPSKTFENTHIKSGGEQKVVLISIGRVKNAEQLKGHDLAARTMGQVAEIKNVRWRVRGINEDDFEKSKKILEDNL  
QSGRLKPTLIPYGTQKEISDDMKMAHVVLMPSEFFGLVGLEAIAADHNEHKTKLQDIMPKEEEELVQLVSQA  
KKKLTALGEKPQYIREAVASIETNMQTVIQDVIEATTESIVRLEESMKTQRDALIKQIEKQCRDNISVLSDDRST  
IERDRSLLADTIHLSEAAATHKDALEILSCIEALKELGVLSEDEEKVLAPMLTFVPASNASYSLGSVTTDDGKA  
SDVEEEEIECSMGIGNDSETSSEYISERNKKGSIYPKPIQIVRFLKGKEEGEFDQPVAVIATDKGLFVLDQNNH  
RMQIFDMTGKFKRSFPVSLPEETGSKPCSFMSHPTDRMVVLDSLKLAFLTKKGFLKTLKGQLKEPSIVKY  
LEQAQGIIVGVTDKNELVLFDLKGDFARICLSTECEVPSKQPIGPYFSVANNMTKQLLFYRADFSLAYVIGGW  
GQGGGRFKGPIGRGYTIDSHDHFVVPDQGSRLQYFDRFCNFVGLIADRDDGLNPSCVTRTPRGLVVSDVGKN  
CIFFFNYLPEHNKTCCVII

### **XP\_002229629**

MPTRDPTASVVVGLCSYTRGYDLNWRLHASLRRVCSPDATATKNCAMYMLLSG**DVETNPGPCKLRI**PAPLGGKQ  
RDTHLRGYYDLRVDWSDLAVCCDNCDICYHSSPVVMS PDTGNSLPRDETDLSPAAANLSNKTPPDSSHSTDSCSQD  
SVTAGILPPKSHNWRTLVVNADGIIGKVAQFANLVSYTKPDAILMCETKLGPHHSAEFMPPGYSPPVRRDRASG  
AGGVLIAVRDCYTAVEVEQPSEI

### **XP\_002234262**

MASPKESSLILFSFVLLSIAACILVNNGHSNMPGSNLHLWRKLGMPTRDPTASVVVGLCSYTRGYELNLSLHASL  
RRVCSPDATTTKNCAMYMLL~~CGD~~**DVEPNPGPCKLHMPAPLGGKQRDDHLCGYCDL**RVDWSDLAVCCDNCDIWHYHR  
PCHGMTTEFGNVHDVHWTCFKCNSSNSDNNTFHSFELDSNSYLYHDTSDPSSVSPSSFRPSYHSSPVVSPD  
TGNNAPRDGTEPDSSHSTDSCQDSVTAGLLPPKSHNWRTLVVNADGIIGKVAQFANLVSYTKPDAILMCETKLGS  
HHSAEFMPPGYSPPVRRDRASGAGGVLIAVRDCYTAVEVEQPSEIKAESIWIKVSLRNKKSLYLGSYYRTPSGK  
HSEQLDELEKSISIHIRNLTKNNPGSTIIGGDFNLGDVDWETESIPKNSRERVPGERLISIFNNNNRVSQMREPTR  
FGRLLDLFVTSKPSLVKTMNSLPGIAYDNIRELFGAFARELDTNINQHTVEESWSALKNKIHDLISSHVPSKTT  
KRHHLLWMNGELKR