

Figure S1. Sequence alignments of subunits in Pol I and their homologs in Pol II.

The alignment is taken from Kuhn et al. 2007. Regions previously assigned to be of conserved fold are underlined and marked green. Additional regions of conserved fold derived from crosslinking and MS analysis are depicted in cyan (A-B cross-links) and yellow (B*-B cross-links). All lysin residues involved in crosslinks are bold. Additional regions of conserved fold might exist but cannot be predicted with certainty.

A190-Rpb1 edited by hand according to 3D structure, EM density and crosslinking-MS

```

A190      ---MDISKVPGSEITSVDFGILTAKAIRNLSAKQITNPTVLDNLG-HPVSGGLYDLALGA 56
Rpb1      MVGQYSSAPLRTVKEVQFGLFSPPEEVRAISVAKIRFPETMDETQTRAKIGGLNDPRLGS 60
          *          * * *          * * *          * * *          * * *          * * *          * * *          * * *          * * *

A190      FLRNL-CSTCGLDEKFCFGHQGHIELVPVCYNPLFFNQLYIYLRASCLFCHHFRLKSVE- 114
Rpb1      IDRNLKCQTCQEGMNECPGHFGHIDLAKPVFHVGFIAKIKKVCCEVCMHCGKLLLDHNE 120
          * * * * *          * * * * *          *          *          *          *          *          *          *

A190      VHRACKLRLLQYGLIDESYKLEITLGSLNSSMYTDEAIEDNEDEMDGEGSKQSKDISS 175
Rpb1      LMRQALAIKDSKKRFAAIWTLCKTKMVCE----- 149
          * * *

A190      TLLNELKSRKRSEYVDMAIAKALSDGRRTTERGSFTATVNDERKKLVHEFHKKLLSRGKCDN 235
Rpb1      -----TDVPE--DDPTQLVSRGG 166
          *

A190      CGMFSPKFRKDGFTKIFETALNEKQITNNRVKGFIRQDMIKKQQKAKLDGSNEASANDE 295
Rpb1      CGNTQPTIRKDGLKLVGS----- 184
          * * * * *

A190      ESFDVGRNPTRPKTGSTYILSTEVKNILDVFRKEQCVLQYVFHSRPNLSRKLKADSF 355
Rpb1      -----WKKDRATGDAEPELRVLSTEEILNIFKHISVKDFTSLGFNEVFSRPEWM 234
          *

A190      FMDVLVVPTRFRLPSKLGEVHENSQNQLLSKVLTTSLLRDLNDDSKLQKDKVSLED 415
Rpb1      ILTCLPVPPPVPRPSISFNESQR---EDDLTFKLADILKANISLETLEHGNAP----- 285
          * * * * *          *          *          *          *          *

A190      RRVIFSRLMNAFVTIQNDVNAFDSTKAQQ-RTSGKVPIPGVKQALEKEGLFRKHMMGK 475
Rpb1      --HHAIEEAESLLQFHVATYMDNDIAGQQALQKSGRPVKSIRARLKGKEGRIRGNLMGKR 344
          *          *          *          * * * * *

A190      VNYAARSVISPDPNIETNEIGVPPVFAVKLTYPEVTAYNIAELRQAVIGPDKWPGATQ 535
Rpb1      VDFSARTVISGDPNLELDQVGVPKSIAKTLYPEVVTPYNIDRLTQLVRNGPNEHPGAKY 404
          *          * * * * *          *          * * * * *          *          * * * * *          * * *

A190      IQNEDGSLVSLIGMSVEQRKALANQLLTPSSNVSHTLNKVYRHIKNRDVVLMNROPTL 595
Rpb1      VIRDSGDRIDLR-----YSKRAGDIQLQYGWKVERHIMDNDPVLFNRQPSL 450
          *          *          * * * * *          * * * * *          * * * * *          *

A190      HKASMMGHKVRVLPNEKTLRLHYANTGAYNADFDGEMNMHFPQNENARAEALNLANTDS 655
Rpb1      HKMSMMAHRVKVIPYS-TFRLNLSVTSPYNADFDGEMNLHVPQSEETRAELSQLCAVL 509
          * * * * *          * * *          * * * * * * * * * * * * * * * * *

A190      QYLTPTSGSPVRGLIQDHISAGVWLTSKDSFFTREQYQQYIGCIRPEDGHTTRSKIVTL 715
Rpb1      QIVSPQSNKPCMGIVQDTLCGIRKLTLRDTFIELDQVLNMLYWVPDWDG-----VIP 561
          *          * * * * *          * * *          *          *          *

A190      PPTIFKPYPLWTGKQIITVLLNVTPPDMPGINLISKNKIKNEYWGKGSLENEVLFKDGA 775
Rpb1      TPAIKPKPLWSGQILSVAIP-----NGIHLQRFDEGTLLSPKDNGMLIDGQ 611
          * * * * *          * * *          * * * * *          * * * * *

```


A135-Rpb2 edited by hand according to 3D structure, EM density and crosslinking-MS

A135 MSKVIKPPGQARTADFRTLERESRFINPPKDKSAFFLLQEAVQPHIGSFNALTEGFDGGI 60
Rpb2 MSDLANSE--KYYDEDPYGFEDESAPITAEDSWAVISAFFREKGLVSQLDSFNQFVDYTL 59
** * * * * *

A135 LNLGVKDIGEKVIFDQKPLNSEDEISNSGYLGNKLSVSVQVSIKPMSTDGVSSAVERK 120
Rpb2 QDIICEDS--TLILEQLAQHTTE---SDNISRKYEISFGKIYVTKPMVNE--SDGVTHA 111
* * * * * * * * * *

A135 VYPSESQRRLTSYRCKLLKLVN---GEENLFVDRD-----CGGLP 162
Rpb2 LYPQEARLRLNLTYSGLFVDVKKRYEADVPGRELKYELIAEESDDESGKVFIRGLP 171
** * * * * * * * * *

A135 VMLQSNRCHLNKMSPYELVQHKESDEIGGYFIVNGIEKLRMLIVQRRNHPMAITRPSF 222
Rpb2 IMLRSKNCYLSEATESDLYKLEKCPFDMGGYFIINGSEKVLIAQERSAGNIVQVFKKAAP 231
** * * * * * * * * *

A135 ANRCASYSHYGIQIRSVRPDQTSQTNVHLNLDGQVIFRFSWRKNEYLVVVMILKALCH 282
Rpb2 SPISHVAEIRSALEKGSRFISTLQVKLYGREGSSARTIKATLPYIKQDIPVIVIFRALGI 291
* * * * * * * * *

A135 TSDREIFDGIICNDVKDSFLTDRLLELLRQFKKRYPHLQNRTOVLQYLGDKFRVVFQASP 342
Rpb2 IPDGEILEHIC-YDVNDWQMLEMLKPCVEDG---FVIQDRETALDFIGR--RGTALGIK 344
* * * * * * * * *

A135 DQSDLEVGQEVLDRIVLVHLGKDG--SQDKFRMLLFMIRKLYSLVAGECSFDNPDAHQHQ 400
Rpb2 KEKRIQYAKDILQKEFLPHITQLEGFESRKAFFLGYMINRLLCALDRKQDDRDHFHGK 404
* * * * * * * * *

A135 EVLLGGFLYGMILKEKIDEYLNIIAQVRMDINRGMALNFKKRYMSRVLMRVNE NIGSK 460
Rpb2 RLDLAGPLLAQLFKTLFKKLTKDI FRYMQRTVEEAHDFNMK-----LAINAKTITSG 456
* * * * * * * * *

A135 MQYFLSTGNLVSQSGLDLQVSGYTVVAEKINFYRFISHFRMVHRGSFFAQLKTTVVRKL 520
Rpb2 LKYALATGNWGEQK-KAMSSRAGVSVQLNRYTYSSTLSHLRRTN-TPIGRDGKLAQPRQL 514
* * * * * * * * *

A135 LPESWGFLCPVHTPDGSPCGLLNHFHAKKRISTQQSDVSRIPSILYSLGVA PASHTFAAG 580
Rpb2 HNTHWGLVCPAETPEGQACGLVKNLSLMSCSISVG-TDPMPIITFLSEWGMPELEDYVPHQ 573
* * * * * * * * *

A135 -PSLCCVQIDGKIIGWVSHECGKIADTLRYKVEGKTPGLPIDLEIG---YVPPSTRGQ- 636
Rpb2 SPDATRVFVNGVWHGV--HRNPARLMETLRLRRKGDINPEVSMIRDIREKELKIFTDAGRV 633
* * * * * * * *

A135 YPGLYLFGG-----HSRMLRPVRYLPLDK-----EDIV 662
Rpb2 YRPLFIVEDDESLGHKELKVRKGHIAKLMATEYQDIEGGFEDVEEYTWSSLLNEGLVEYI 693
* * * *

A135 GPFEQVYMNIAVTPQEIQ-----NNVHTHVEFTPTNLSILA 701
Rpb2 DAEESILIAMQPEDLEPAEANEENDLDVPAKRIRVSHHATTFTHCEIHPSMILGVAA 753
* ** * * * * *

A135 NLTPFSDFNQSPRNMYYCQMGKQTMGTPGVALCHRSDNKLYRLQTGQTPIVKANLYDDYG 761
Rpb2 SIIIPFDHNQSPRNTYQSAMGQAMGVFLTNYNVRMDTMANILYPOKPLGTTTRAMEYLK 813
** * * * * * * * * *

A135 MDNFPNGFNAVAVISYTYGDMDDAMIINKSADERGFYGTMYKTEK-VDLALNRNRGDP 820
Rpb2 FRELPAGQNAIVAIAACYSYGNQEDSMIMNQSSIDRGLFRSLFFRSYMDQEKKYGMSITET 873
* * * * * * * * *

A135 ITQHFQFGNDEWPKEWLKLEDEDGLPYIGTYVEEGDPICAYFDDT-----LNKTKIKT 873
Rpb2 FEKQRTNTLRMKHGTDKLDDGLIAPGVRVSGEDVIIGKTPISPDEEELGQRTAYHS 933
* * * * * * * *

A135 YHSSEPAVIEEVNLIGDESNIKFOE--LQTVSIKYRIRRTPOIGDKFSSRHGQKGVCSRK 930
Rpb2 KRDASTPLRSTENGIVDQVLVTTNQDGLKFKVVRVTRTKIPQIGDKFASRHGQKGTIGIT 993
* * * * * * * * *

A135 WPTIDMPFSETGIQPDIIINPHAFPSRMTIGMFVESLAGKAGALHGIAQDSTPWIFNEDD 990
Rpb2 YRREDMPFTAEGIVPDLIINPHAFPSRMTVAHLIECLLSKVAALSGNEGASPF----D 1049
**** ** * * * * * * * * *

A135 TPADYFGEQLAKAGNYHGNPEMYSYGATGEELRADIYVGVVYVYQRLRHVMNDKQVVRSTG 1050

Rpb2 ITVEGISKLLREHGYQSRGFVEMYNGHTGKKLMAQIFFGPTYQRLRHMVDDKIHARARG 1109
 * ** *

A135 **EVNSLTMQPVKGRKRHGGIRVGEMERDALIGHGTSFLLQDRLLNSSDYTOASVCRECGSI** 1110
 Rpb2 PMQVLTTRQPVEGRSRDGLRFGEMERDCMIAHGAASFLEKRLMEASDAFRVHICGICGLM 1169
 * ** ** *

A135 **LTTQ**QSVPRIGSIS**TVCCRRCSM**RFEDAK**KLLTK**SEEDGE**KI**FIDDSQIWEDGQGNKFVGG 1170
 Rpb2 TVIAKLN-----HNQFECKGCDN-----K 1188
 * *

A135 **NETTVAIPFVLKYLDSELSAMGIRLRYN**VEPK--- 1203
 Rpb2 IDIYQIHIPYAAKLLFQELMAMNITPRLYTDRSRDF 1224
 * * * * * * * * * *

AC40-Rpb3 edited by hand according to 3D structure, EM density and crosslinking-MS

```

AC40      MSNIVGIEYNRVTNTTSTDFPGFSKDAENEWNVEKFKKDFE VNISSLDAREANFDLINID 60
Rpb3      -----MSEEGPQVKIREASKDNVDFILSNVD 26
              * * * * *

AC40      TSIANAFRRIMISEVPSVAEYVYFFNNTSVIQDEVLAHRIGLVPIK-VDPDMLTWVDSN 119
Rpb3      LAMANSLRRVMIAEIPTLAIDSVEVETNTTVLADEFIAHRLGLIPLQSM DIEQLEYSRDC 86
              * * * * * * * * * * * * * * * * *

AC40      LPDDEKFTDENTIVLSLVKCTRNPDA PKGSTDPKELYNNAHVYARDLKFFPQGRQSTTF 179
Rpb3      FCED--HCDKCSVVLTLQAFGESE-----STTNVYSKDLVIVSNLMGRNIG 130
              * * * * * * * * * *

AC40      ADCPVVPADPDLLAKLRPGQEISLKAHCILGIGGDHAKFSPVSTASYRLLEQINILQPI 239
Rpb3      HPIIQDKEGNGVLI CKLRKGQELKLT CVAKKGI AKEHAKWGPAAAIEFEYDPWNK LKH-- 188
              * * * * * * * * * * * * * * *

AC40      KGESARRFQKCFPPGVIGIDEGSDEAYV KDARKDTVSREVLRYEEFADK---V KLGRVRN 296
Rpb3      -----TDYWYEQDSAKEW PQSKNCEYEDPPNEGDPFDYKAQAD 226
              * * * * * * * *

AC40      HFIFNVESAGAMTPEEIFFKSVRILKNKAEYLKNCPI TQ----- 335
Rpb3      TFYMNVESVGSIPVDQVVVRGIDTLQKKVASIL-LALTQMDQDKVNFASGDNNTASNMLG 282
              * * * * * * * * * *

AC40      ----- 356
Rpb3      SNEDVMMTGAEQDPYSNASQMNGTGGGGYD NAW 318

```

AC19-Rpb11 edited by hand according to 3D structure, EM density and crosslinking-MS

AC19 MTEDIEQ**KK**TATEVTPQEPKHIQEEEEQDVMGTGDEEQEEEPDREKI**KLLTQATSE**DGTS 60
Rpb11 -----MNAPDRFELFLLGEGESKLLKIDPDTKAPNA 30
* *

AC19 **ASFQIVEEDHTLGNALRYVIMKNPDVEFCGYSIPHPSENLLNIRIQTYGETTAVDALQK**G 120
Rpb11 VVITFEKEDHTLGNLIRAELLNDRKVLFAAYKVEHPFFARFKLRIQTTEGYDPKDALKNA 90
***** * * * * * * * *

AC19 **LKDLMDLCDVVE**SK**FEKIK**SM----- 142
Rpb11 CNSIINKLGALKTNFETEWNLQTLAADDAF 120
*

A12.2-Rpb9 edited by hand according to 3D structure, EM density and crosslinking-MS

A12.2 MSVVGSLI **FCLDCGDLI** ENPNAVLG---SNV**EC SOCK** AIYPKSQFSNL**KVV**TTTADDAFPSSLRAK**KSVVKT**SL 71
Rpb9 ---MTTFRFCRDCNNMLY-PREDKENRLLFECRTCSYVEEAGS-----PLVYRHELITNIGETAGVVQ 60
* * * * * * * * * * * * * * *

A12.2 **KKNELKDGATIK** **KCPQCG** NEEMNYHTL**QLR**SADEGATV**FYCTSCGYK** RTNN----- 125
Rpb9 DIGSDPTLPRSDRECPKCHSRENVFQSQRRKDTSMVLFVCLSCSHIFTSQKPKRTQFS 122
* * * * * * * * * * * * * * *

A12.2 C-terminus -----RAK**KSVVKTSLK**KNE---L**KDGATIK****KCPQCG**NEEMNYHTL**QLR**SADEGATV**FY** 52
TFIIS 3rd domain PAPLKQKIEEIAKQONLYNAQGATIERSVTDRTCGKCKEKKVSYQ**LQ**TRSADEPL**TTFC** 60
* * * * * * * * * * * * * * *

A12.2 C-terminus TCTSCGY**K**FRTNN 65
TFIIS 3rd domain TCEACGNRWKFS- 72
* * * *