Supplementary information

The crystal structure of γ -tubulin complex protein GCP4 provides insight into microtubule nucleation

Valérie Guillet^{1,2,6}, Martine Knibiehler^{3,6}, Lynn Gregory-Pauron^{1,2,6}, Marie-Hélène Rémy³, Cécile Chemin³, Brigitte Raynaud-Messina³, Cécile Bon^{1,2}, Justin M. Kollman^{4,5}, David A. Agard^{4,5}, Andreas Merdes³ & Lionel Mourey^{1,2}

¹Institut de Pharmacologie et de Biologie Structurale (IPBS), Centre National de la Recherche Scientifique (CNRS), Toulouse, France. ²Université de Toulouse, Université Paul Sabatier, IPBS, Toulouse, France. ³Centre de Recherche en Pharmacologie-Santé (CRPS), UMR 2587 CNRS-Pierre Fabre, Toulouse, France. ⁴Department of Biochemistry and Biophysics, Howard Hughes Medical Institute, University of California, San Francisco, California, USA. ⁵Keck Advanced Microscopy Center, University of California, San Francisco, California, USA. ⁶These authors contributed equally to this work. Correspondence should be addressed to A.M. (amerdes@cict.fr) or L.M. (lionel.mourey@ipbs.fr).

						α1 00000000	2	β1	β2
GCP4_HUMAN GCP4_MOUSE Q64233_XENLA AA163276 DANRE GCP4_DROME GCP4_DROME GCP4_MEDTR A8HP50_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	1 1 1 1 1 1 1	MCRYELLG	MGI GRGLLYWG	(AFESGI)RRKSKA	VPFDESG	MIHELLIAI MIHELLIAI MIHELLIAI MIHELLIAI MIHELLIAI MIHELLIAI MIHELLIAI MIHEILISI MIHEILISI	SGYPG.SI SGYPG.SI SGYPG.SI SGYPG.SI SGYPG.SI SGYPG.SI LGFTG.DL LGFTG.DL SGYTG.DL SGYPS.PL AGNTSGGL	TWN. TWN. TWN. GI. UDE. UDE. VDS. AFN. IEK.	.KRSGLQV KRSGLQV KRNGLQV KRNGLQV REQRKTLGLAF RDNNL SDIDRFGPLLA TEEDIVP NGQILLH
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	207 240 1 251 326	YSSDPLYV LHQGELQI	PDDR		LASQESA MEITEAA VLVTETQ VLVKECE	VVEDLLYVI LVRDILYVI MIHELLLAI VIRETLWLI LVKDVLNVI	VGVDGRYVS FQGIDG.KNJ SGYPG.SIF SGVKKLFIF IGVVSATFS	AQPLAGRQS KMN TWN QLID LCQPAQAF	SRT.FLVD .NTENCYK .KRSGLQV GKVTVRNN VVKRGVHV
GCP4_HUMAN GCP4_MOUSE Q64253 XENLA AA163276 DANRE GCP4 DROME GCP4_MEDTR A8HP50 CHLRE Q4WZU3_ASPFU GFH1_SCHPO	28 28 28 26 32 26 48 56 28	NSDSPLSI SANTPISI LDDEKEEB)E DE SWNLPG]	CTFK CTFK SHCTVT	. SQDFPF . SQDFPF . SQDFPF . SQDIPF . SQDLPF . TVIDQF LAPDISF LAPDISF LAPDISF LAPDISF LAPDISF LAPDISF . PSVQP.	LHPSETSVI LHPSETSVI LHPGETSVI IHPGETSVI IHPCEREII TEPSERDL IDPSDREL ISEADRAVI LSPPEKALI LHPGEREL	LNRLCRIGTI NRLCRIGTI NRLCKIGTI NRLCKIGTI MDIIKIIKI IERLIKIGFY GDLIKIGFY GDLIKGFY TSLARISRI INDIVSLASI	YIRFTEFI YIRFTEFI YIRFTEFI YVRFTEFI YVRELORFA YRELDRFA YRELERFS IFKAVAAFY HAKLRE. KRKRIERYI	QQ EQYTGHVQQQ EQYTGHVQQQ EQYTGHVQQQ EQHTGHVHQQ HSSGRKSDT KKSRNLSWIRSV AKSRNLSWIRS. EREQSPWDPR QTSLIS EIFDEDLIQQ
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	250 277 28 300 386				PN .VEGKAN .SQDFPF .IIVTHI SGA	LDLSIRELY LSRSLRDT2 LHPSETSVI THSCLRSVI SPESISSLI	THRTL PVAAS VRLSELGWI UNRLCRLGTI LEQIAAYGQV LSEVAEYGTO	SYSAVTRFI LHNKIRRYT DYIRFTEFI VVFRLQEFI CYT <u>RLS</u> HFS	EEKSSFE DQRS.LDR EQYTGHVQQQ DEVMGHSSES LQPVLDS
					مومومو			222222	
GCP4_HUMAN GCP4_MOUSE Q642S3_XENLA AAI63276_DANRE GCP4_DROME GCP4_ARATH GCP4_MEDTH A8HPS0_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	70 70 70 67 90 83 112 91 69	TSVHPLEF ENANPLE	DHHPSQ DHHPPQ DHHPSQ EHYSSQ ADELSK SSHPSI	QGQGGLH QGQGGLH QGQVGLH PSQSGLH SLPDSLH QSREKKP . NKEKP HATRSP ICRAVSS NQFL	GIYLRAF GIYLRAF GIYLRAF GYYLLNI SVYRRAI SVYRRAI SVYRRAI AINAHHI LKALRLI	CTGLDSVLC CTGLDSVL CTGLDSLHSML AKGIEMALI ANGIGEILS ANGIVEIL ASGLTEML GGFQRKIL LSKKLFLY	PYROALIDI PYROALIDI PYROALIDI PYROALIDI EYYAEIGRI SVYRSAVIHI AVYSSIIH AVYSSIIH EVYRSAVIHI SVEKAILAE DYRCTVIT	EQEFLGD EQEFLAD EQEFLAD EKEFLGD EKYCLGN EQKLLAE EQLLISE JQELRIE SAYVGGY 11ISSDPY	
GCP2 HUMAN GCP3 HUMAN GCP4 HUMAN GCP5 HUMAN GCP6 HUMAN	285 316 70 342 422	· · · · · · · · · · · · · · · · · · ·	DHHPSQ ML.PGS	GQGGLH SVPKKS	GQVNHAI GLVGQSF GIYLRAF TEAPFRT LYSKGLV	AAAMRTLVI CAALHQELI CTGLDSVL YQAFMWAL FQAFTSGL	KEHLILVSQI REYYRLLSVI OPYRQALLDI KKYFISFKEH RRYLQYYRAG	LEQLHRQ LHSQLQLEDI LEQEFLGD. LAEIEKCI CVLSTPP	GUINE CONTRACTOR CONTR
		α5 <u>0000</u>		000000	<mark>α6</mark> 0000000		200	α7 <u>00000000</u>	0000
GCP4_HUMAN GCP4_MOUSE Q64233_XENLA AA163276_DANRE GCP4_DROME GCP4_ARATH GCP4_MEDTR A8HP50_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	116 116 116 109 144 128 151 137 106	HLSISHVN HLSISHVN HLSISHIN RNSLSYVY TPILATVI MPILATVI APVLSTIN VGSGKFVS	YFL YSL YSL YML EGL QGL GEFSPW SLAQLEAN	DQFQLL DQFQLL DQFHLL EQFQLL NKFFVL NKFFVL NKFFSL VEYMDVL TRRLEWL HLADWKE	FPSVMVV FPSVMVV FPSIMVV FPSVMVV LVFMRNI FPPLYEV LPPLYEI PALHALA WDVTRFV TLSSLLV	VEQII VEQII VEQII VEIII I. VEIII I. IIEII I. IIEII I. ILEII AAQDI LPENI	KSQKIHG KSQKIHG KSQKIHG KSQKIHG HVLNLRG ERDDIRG ERDDIRG ERDDIRG ERSDIRG EVSFNRLFP2 RTSSAHGCT(ISQCLS.	CQILETVY CQILETVY CQILETVY CQILETVY CVLLHNLH GQLLNVLN GELLNLHN ADLINRLA SSALIDYLR VASTFQ	KHSCGGL. KHSCGGL. KHSCGGL. QCEHGD. KRCHCGV. KKCHCGV. KGTRSGL. TESQTGY. KLHNMYM.
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	318 361 116 399 460	LLSLQKLV SLTLRRLI HLSISHVN VDKLAPRI GFLFKKL	NFYI LVWT NYFL LSQL GRQL	QPAMRT YDPKIR DQFQLL KVLHKV RYLAEL	MDIL <mark>ASI</mark> LKTLAAI FPSVMVV FSTGVAE CGVG <mark>AV</mark> I	ATSVI VDHCO VEQII VPPD PGTCGGGPI	DKGECLG QGRKG KSQKIHG FRNVVRA RAAFPTG	.GSTLSLLH GELASAVH CQILETVY SHLLNTLY .VKLLSYLY	DRSFSYTGD. AYTKTGD. KHSCGGL. KAILEYDNVGEA QEALHNCS.
			c	(8			β3		
GCP4 HUMAN GCP4_MOUSE Q642S3 XENLA AAI63276 DANRE GCP4_DROME GCP4_ARATH GCP4_MEDTR A8HPS0_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	166 166 166 159 194 178 205 194 157	PPVRSAI PPVRSAI PPVRSAI IQLEKAI PELRTCI PELRTCI PALQTCI LDLEEIA GSTHLNE	EKILAVC EKILAVC EKILAVC EKILAVC KIIMKP QRLLWNC QRLLWNC QRLLWNC GRLLWHC LHLITV FKKLVLEC	HGVMYK CHGVMYK CHGVMYK CHGVMYK CHGVMYK SHQVMYN SHQVMYN CNQVLMQ AESAWIR CECAFQK	QLSAWMI QLSAWMI QLSAWMI QLAAWMI QLAAWMI QLAAWMU QLASWMV QLASWMV QLASWWI QLSTWLI TWLDELV	HGLL.LDQ HGLL.LDQ HGLL.LDQ FGVI.DDV YGIL.CDV YGIL.EDR HGML.QDP YGNLPIFG QYLVNDSG	HEFFIKQGI HEFFVRQGS SEEFFVRQGS SEEFFVRQGS ISEFFIKRQ IGEFFIKRQ IGEFFISRQ VQEFFIQFR VQEFFIQRD DDSFKCSFI	адррррааа	SSGNVS SSGTLS SSGTLS SSGNLA SAGGAT ATQQPQPQPQPQPQ SS CGDTKD
GCP2 HUMAN GCP3 HUMAN GCP4 HUMAN GCP5 HUMAN GCP6 HUMAN	370 409 166 454 515	. SQAQELC . PYMRSLV . PPVRSAI SEQTVSLI . NEHYPVI	LYLTKA QHILSLV EKILAVC FSLWVE LSLLKTS	ASAPYFE 7SHPVLS CHGVMYK CVRPYLQ SCEPYTR	VLEKWIY FLYRWIY QLSAWML TVDEWIV FIHDWVY	RGII.HDP) DGEL.EDT) HGLL.LDQH HGHL.WDG2 SGVFGRDA)	(SEFMVEEH (HEFFVASDI HEFFIKQGI AREFIIQRNI (GEFMIQVNI		

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2083

								β4 η1 → 000	α 9 00000000
GCP4_HUMAN GCP4_MOUSE Q64233_XENLA AAI63276_DANRE GCP4_DROME GCP4_ARATH GCP4_MEDTR A8HP50_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	216 216 216 202 237 221 269 241 207	AQPEDEEDL AQLEEDEEDL AAFEEEEDDL AAQEEDEEDL DDG EGR QQTQAQGGPQ 	GIGGLTG GIGGLTG GIGGLTG GIGGLSG PTDAVDG DLDHRSS DVENSSS QQATREE	KOLRELO KOLRELO KOLRELO SSFSKSZ QEEVSEH HOEISEH DSAGASJ	DLRLIEE DLRLIEE DLRLIEE DLRLIEE LCSLLSA LARTSVH SGRGADD	E N) E N) E N) E N E N E N D A Q M	MLAPSLKQ MLAPSLKQ MLAPSLKQ MLAPSLQQ NPEDYIWQ SLTDWHSG SLSDWHMG AYREWHAG DDSSSITN PFLTQD	FSLRVEI FSLRVEN FSLRVEN YEVNMSC FHISLDM FHISLDM FQVNPRC FVMQTSI IVNCLVS	LPSYIPVRVAEKIL LPSYIPVRVAEKIL MLPSYIPVRVAEKIL JLPGFFSIVLAEKVL MLPDYIPMRLGESIL LPDYIPMRLGESIL LPLYVSCDLADTIL LPKFVSLHTASSIL JIRTCMLNAKKCFTD
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	414 453 216 499 559	AQPEEDEEDL NVPVNH EYLSF	L TV GIGGLTG RDFWYAT RDK	RKERIQI KTDRLWI KQLRELÇ YTLYSVS	D D DLRLIEE EKTENEE 	Y EN KMSDNA S	NDKYWDQR K MLAPSLKQ SASSGSDQ YWTHGYVL	YTIVQQQ YTLRKSM FSLRVEI GPSSRQH ISKEVED	2IPSFLQ.KMADKIL IIPSFMTMDQSRKVL IIPSYIPVRVAEKIL ITMVSFLKPVLKQII OCVPVFLKHIAHDIY
GCP4_HUMAN GCP4_MOUSE Q64253_XENLA AA163276_DANRE GCP4_DROME GCP4_DROME GCP4_MEDTR A8HP50_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	278 278 278 254 292 276 331 272 242	FVGESVQMFE FVGESVQMFE FVGESVQMFE FVGESVQMFE FVGEVQMFE FVGEVQMFE	NQNVNLT NQNVNLT NQNVNMS NHNQSPS MGRNVKV NPSPAFQ NPSPSF RPGQGGG AK RLVYEMP	KNKTDPI FQKDKSI LSQDDV LPLT.	.AAKL VQVTMRGS FQEPKRF	AELDSD QRIRGF PKIHGF	D MHSDFPET EGRFNFQR A	RKGS RTGS RTGS RAGS RAGS 	Q SILKNQ SILKSE SILKSE SIGAGAELLPQSEADK SIGAGAELLPQDMMK SIMGFDSGAPVTLQ
GCP2 HUMAN GCP3 HUMAN GCP4 HUMAN GCP5 HUMAN GCP6 HUMAN	452 485 278 561 597	ST <mark>GKYLNVVR</mark> LI <mark>GKSINFLH</mark> FV <mark>GESVQMFE</mark> MAGKSMQLLK VC <mark>GKTINLLK</mark>	ECGHDVT QVCHDQT NQNVNLT NLQCAE. LCCPRHY	CP	STTC VPRISVI	 QA FS	· · · · · · · · · · · ·	RKGS GARD LEEI	
		<mark>α10</mark> <u>0000000000</u>	٥٩	معامد	α11 20000000	00000			
GCP4_HUMAN GCP4_MOUSE Q642S3_XENLA AAI63276_DANRE GCP4_DROME GCP4_ARATH GCP4_MEDTR A8HPS0_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	304 304 304 300 358 341 370 302 263	EDTFAAELHR EDTFAAELHR EDTFAAELHR EDMFAAELHR FFKMVVDLSN IEAMLKDLKE IENMLLDLKE FGQAFTALQS GEHIAHLASL KSSVTKLA	LKQQPLF LKQQPLF LKQQPLF ESSEF SSEF EPEL GSPI DEMFTLV	S.LVDFE S.LVDFE S.LVDFE S.LVDFE N.VFRLF H.KRSFE S.APRL S.APRL S.ASKLS SGEKIIE	OVVDRIR OVVDRIR SVLDRIR NLVDGIR CTVDSVR CAVDSIQ MLVEGLR NAVDSIR OLASFHQ	STVAE STVAE STVAE STVAE NYVSA AIAAS AIAAS SHVAR LSLSQ IFDLV			
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	471 507 304 592 639	AKEIIYTLKE IAVTKSAESF EDTFAAELHF LYTLFLESVQ DCAVYVGRME	QDAADLF LKQQPLF SRLRHGE RVARHSS	RAYVI TDENAF(S.LVDFI DSTPQVI VSKEEKI	EQ.IEKAF QGKIDAAY EQVVDRIR LTEQQATK ELR <mark>ME</mark> IAK	NY <mark>A</mark> SK FETSK STVAE ENLMK. QELIA.	77-resi · 827-resi	due-long	insertion
		2222222	α12 202020	2 QQQQQ •	معمع	α13 0000000	معمعه		α14 2020202020200
GCP4_HUMAN GCP4_MOUSE Q64253_XENLA AAI63276_DANRE GCP4_DROME GCP4_ARATH GCP4_MEDTR A8HPS0_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	339 339 339 332 390 373 402 334 296	HLWKLMVEES HLWKLMVEES HLWKLMVEES KLSEIAVNEV HLWQLVVVRA HLWQLVVVRA LLWRLLVGRA SVLSKLPLP KMLEMVMLLE	DILGQLK DILGQLK DILGQLK DILGQLK DILGQLK DILGQLK DILG GLL F A C L L C G L L P C G L L P C G L L P C G H L K C G H K C C L K C C C L K C C C L K C C C L K C C C C	IIKDFYI IIKDFYI IIKDFFI LIKDFFI ALKDYFI ALKDYFI ALKDYFI VLHDFLI L <u>SEMFRI</u>	LGRGELF LGRGELF LGRGELY LGRGEFY LEKGDFF LAKGDFF LAKGDFF LGRGDFW	QAFIDT QAFIDT QAFIDV QVFIDL LEFCSQ QCFLEE QAFLAE TALVSH SS.NDV	AQHMLKTP AQHMLKTP AQNMLKTP AQNMLKTP SRQLMRLP SRQLMRLP SRQLMRLP ARPLMAAP ADARLSER QGVVIDDI	PT.AV PT.AV PT.AV PS.AV PRQST PRQST PRPGS HRIPE LPVVL	. TEHDVNVAFQQSA TEHDVNVAFQQSA . TEHDVNVAFQQSA . TEHDVNVAFQQSA . FKNVTRSFELAAT . GESDLMVPFQLAS . AEADLMVPFQLAS . VDADLALPFTRSA . LLAGRVRAQTELE . RKTISLLDLATEY
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	498 543 339 705 1502	VLLDFLMEEK YLLDVLNKKY HLWKLMVEES NLMQTLKKDY AAVDYFFVEL	ELVAHLR SLLDHMQ DLLGQLK RLVEYLQ HLEAHYE	SIKRYFI AMRRYLI IIKDFY AMRNFFI ALRHFLI	MDQGDFF LGQGDFI LGRGELF MEGGDTM MEDGEFA	VHFMDL RHLMDL QAFIDT YDFYTS QSLSDL	AEEELRKP LKPELVRP AQHMLKTP IFDKIREK LFEKLGAG	VE.DI AT.TL PT.AV ET.WQ QT.PGEI	.TPPRLEALLELAL .YQHNLTGILETAV .TEHDVNVAFQQSA .NVSFLNVQLQEAV .LNPLVLNSVLSKAL
		٥٥		<mark>η2</mark> .ደደ ደ■	β5				
GCP4_HUMAN GCP4_MOUSE Q642S3_XENLA AA163276_DANRE GCP4_DROME GCP4_ARATH GCP4_MEDTR A8HPS0_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	401 401 401 394 453 436 465 397 357	HKVLLDDD. HKVLLDDD. HKVLLDDD. HKVLLDDD. TKTIAEED. LKTIGEED. AKSSAEGD. GLVVKDGD. GECVSFLK.		NLLP.LI NLLP.LI NLLP.LI NLLP.LI XYFS.K YFS.K PLLA.AI VVHA.L2 LVSIGFI	HLTIEYH HLTIEYH NLTIDYH HLTIDYH SLRMPSF YSLRMPSY YQLRY.LG HTWAELY	GKEHK I GKDHKA GKEHK I GKESK I GVTVRS GITVKP GKAEA I SLQNEE	DATQAREG DATQPREV DTSQPREG EGSGNREG SQADMVR. SQADMVR. SLLNVPKA EAAFTVKH DPVDESLD	LARDLVF	TSA RLSITDSKNRHTVAT
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP5_HUMAN	560 605 401 767 1567	RMSTANTDPF RATNAQFD HKVLLDDD GQRY QCSLHGDT	KDDLKID	LMPHDL] EILRRLI NLLP.LI PEDSSRI PHAS.NI	TQLLRVL VRLL HLTIEYH SISFENV SLALKYL	AIETKQI GKEHK.I DTA PEVFA.I	EKAMAHAD DATQAREG KKK. PNAP	· · · · · · · · · · · · · · · · · · ·	

			β6	η3 000 000000	α15 000000000000000000000000000000000000
GCP4_HUMAN GCP4_MOUSE Q64253_XENLA AA163276_DANRE GCP4_DROME GCP4_ARATH GCP4_MEDTR A8HP50_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	435 436 435 408 487 474 498 453 375	PSRETSPREAPASGWAA PSRETSPREAPSSGWAA PFRDMSPREAPSSGWAA CORSTSPREAPTSGWAA CORSTSEPDDTSDFNFI SKVSLTGK.ANLTSDTSVDGWDA AADGISGASISNASSEMSVDGWDA CORSTSAHSQVPELDRAWDI HSDEIGSDLVAELSDVSFEDLLFI MIRSENRKIDVEGLM	ALGISYKVQ. W ALGLSYKVQ. W ALGLSYKVQ. W ALGLAYKVQ. W GLSLKYEYEW ALALEYSVD. W IALEYSVD. W IALEYSIE. W ZVLEVGLE. W MQLLLCLK. W	PLHILFTPAVIEK PLHILFTPAVIEK PLHILFTPAVIEK FLHILFTPAVIEK FLHILFTPAVIEK PLNLFSPTTIER PLILFFTQEVISK PLHLFFTQEVISK FLGLLGREQLRR FLGLLFLSTSDIAV PFNLFMDKNALNV	YNVVFKYLLSVRRVQAEL YNVVFKYLLSVRRVQAEL YNVVFKYLLSVRRVQSEL YNVVFRYLLSVRRVQSEL YNNIFRFLLIIRTYQYEI YLKVFQYLIRLKRTQMEL YLKVFQYLLRLKRTQMEL YNQLFALLRLRRMQNQL YSRIHSYLLGIRRAQIRI YSDLWILLGSLHLSITKI
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	605 626 435 791 1597	PT. ELALSGLEA 	AFSFDYIVKW VFSLDYHVDG AlglsykvQW GltlsykvPW ClelrykvDW	PLSLIINRKALTR PIATVFTRECMSH PLHILFTPAVLEK PVDIVISLECQKI PLNIVITEGCLSK	YQMLERHMFYCKHVERQL Ylrvfnflwrakrmeyil YnvvfkyllsvrrvQael YnQvfllllqirwarysl YSGVFSfllqlklmmwal
GCP4_HUMAN GCP4_MOUSE Q64253_XENLA AAI63276_DANRE GCP4_DROME GCP4_ARATH GCP4_ARATH GCP4_MEDTR A8HPS0_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	492 493 492 494 467 549 538 555 519 430	14 QHCWALQMQ.RKHLKS QHCWALQMQ.RKHLKS QHCWALQMQ.RKHLKS QHCWALQMQ.RKHLKS QHCWALQMQ.RKHLKS QHCWALQMQ.RKHLKS QHCWALQMQ.RKHLKS QHCWALQMQ.RKHLKS QHCWALQMQ.RKHLKS QRVWAKQTW.RAK.SAKD EKSWASVMH.QDHES EKKUASVMH.QDHSIF DDAWKDLRV.MDREMQ NDLWKHSFL.RKSYPSPWGPI KSLFYERKEGYTAK	AQHRK.D AKNKKSD PRSNTPFGQSKL	QQ NQT NKT NKT NKT 	al6 DAIKWRLRNHMAFLVDNL DAVKWRLRNHMAFLVDNL DAVKWRLRNHMAFLVDNL DAVKWRLRNHMAFLVDNL NNKIITLRNYLMFFLNNM IRPMWRVREHMAFLIRNL FRSMWRVREHMAFLIRNL FRSMWRVREHMAFLIRNL GNRPVWATSSAALFVLSE ISQPWTQLWVTLWFLSSL
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	656 677 492 839 1642	CSVWISNKT.AKQ.HSLHS TDIRKGHMCNAKLLRN QHCWALQMQ.RKH.LKS DVLLFGELVSTAEKPRLKE KDVCFHLKRTALLSHMAGS	GLIHEQDTV		FAGAFT <mark>LRQRMLNFVQNI</mark> LHQCHILASEMVHFIHQM DAIKWRLRNHMAFLVDNL IHRMFLLRVKLMHFVNSL FRQLQL <u>FK</u> HEMQHFVKVI
GCP4_HUMAN GCP4_MOUSE Q64253_XENLA AAI63276_DANRE GCP4_DROME GCP4_ARATH GCP4_MEDTR A8HPS0_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	528 529 528 530 504 601 591 599 580 462	QYYLQVDVLESQFSQLLHQIN QYYLQVDVLESQFSQLLHQIN QYYLQVDVLESQFSQLLQQIN QYYLQVDVLESQFSQLLQQIN QYYLQVDVLESQFSQLLQQIN QYYLQVDVLESQFGILMNVIK QFYIQVDVLESQWKVLQTHIH QFYIQVDVIESQWKVLQTHIH QFYIQVDVIESQWNILQSHIQ GSFFQGEVVNGCWQHFREWIQGG QYYAYECVIKPSYCKLRESLT	JLATATSSTSSR	PETSASLKSRYSD	ST ST ST ST ST ST ST ST ST ST ST ST ST S
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	694 718 528 898 1681	QYYMMFEVMEPTWHILEKNLK QYYITFEVLECSWDELWNKVQ QYYLQVDVLESQFSQLLHQIN HNYIMTRILHSTGLEFQHQVE QGYIANQILHVTWCEFRARLA		· · · · · · · · · · · · · · · · · · ·	
GCP4_HUMAN GCP4_MOUSE Q64253_XENLA AAI63276_DANRE GCP4_DROME GCP4_RATH GCP4_MEDTR A8HPS0_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	551 552 551 527 624 614 622 646 488	RD FE SIRLAHD HFLSNLLAQSFI RD FE SIRLAHD HFLSNLLAQSFI RD FE SIRLAHD HFLSNLLAQSFI RD FE SIRLAHD HFLSNLLAQSFI SD FE VIQAHTVFLANVLSHCFI QD FTELVGFHQEYLSALISQSFI HD FTELVGFHQEYLSALISQFFI QD FSELERAHALFLNTLMVQSFI RLDPETLTVAHHRYLFSLVQSLFI QKLRMQDCSQLHALTLTYAKKML	L LNESETQLNVTG D T	CIP CONTRACTORNEL CAN BE CONTRACTORNEL CAN BE CONTRACTORNEL CONTR	LCHSFCSLVSQNLGP LCHSFCSLVSQNLGP LCHSFCSLVSQNLGP LCHNFCSLVSQNLGS LFGICEKFAHMTQT. LCLQFCWNIENQESN LCLQFCWNIENQENN LCLQFCWNIENQENN LCLQFCWNISRACCD LIDHFIALIIRLESV QEV <u>SK</u> VDCQLAAHVS
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	717 741 551 921 1704	SNIDDVLGHHTGFLDTCLKDCML QDLDHIIAAHEVFLDTIISRCLL RDFESIRLAHDHFLSNLLAQSFI KDLDQLIKIHYRYLSTIHDRCLL GDLEEIQRAHAEYNHKAVFRGLL	5	NPELLKVFSKLMS LNQLRAVFDQIIE LKPVFHCLNEILD EKVSFVKEAIMKV EKAAPVMNVIHSI	VCVMFTNCMQKFTQS LQNAQDAIYRAALEE LCHSFCSLVSQNLGP LNLALMFADGWQAGL FSLVLKFRSQLISQAWGP
GCP4_HUMAN GCP4_MOUSE Q64253_XENLA AAI63276_DANRE GCP4_DROME GCP4_ARATH GCP4_MEDTR A8HPS0_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	603 604 603 605 589 676 666 674 699 549	LD. LD. LD. KD. PN. SN. PA. QK. HS.	ERGAQU ERGAQU ERGAQU ERGAQU PSDDLEDEV SELE RLDFAAV NMDLETDEG IECIRQK	C20 C20CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	LFKILSSVRNHQ LFKILSSVRNHQ LFKILSSVRNHQ LFKILSSVRNHQ LFKILSSVRNHQ LIQLLVDVKSAS YTILRSKLAGS YTILRSSKLAGS LFGSISSQRMVESF LFGSISSQRMVESF LIVQLNL
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	769 799 603 973 1759	MKLDGELGGQTLEHSTVLGLPAGA LQRRLQFEEKKKQREIEGQWGV LD	AEERARKELARK /TAAEEEEENKR ERGAAQL WRMESIEKME ?RGAEHPNFALM	HLAEHADTVQLVS IGEFKESIPKMCS SILVKGFSRQSSL SDFKNCHMFLVTI QQSYNTFKYYSHF	GFEATINKFDKN QLRILTHFYQGI LFKILSSVRNHQ INKAVCRGS LF <mark>KVV</mark> TKLVNR

		α21
		909099999999
GCP4 HUMAN GCP4-MOUSE Q64253 XENLA AAI53276 DANRE GCP4 DROME GCP4 DROME GCP4-MEDTR A8HP50_CHLRE Q4WZU3-ASPFU GFH1_SCHP0	637 638 637 639 625 708 698 710 741	INSDLAQLLLRLDYN INSDLAQLLLRLDYN INSDLAQLLLRLDYN INSDLAQLLLRLDYN CLGPLSQLLLRLDYN ORAPFLSRFLLRLNFN QRTPFLRRFLLRLNFN DVVARLRDIDDSRSGEGRKMLDLLKHARATPLADRKNGSGVNAQICQYLPRKTAGVDQLLMKLDF.
GCP2_HUMAN GCP3_HUMAN GCP4_HUMAN GCP5_HUMAN GCP6_HUMAN	829 856 637 1007 1798	FSAHLLDLLARLSIYSTSDCEHGMASVISRLDFN VQQFLVLLTTSSD ESLRFLSFRLDFN INSDLAQLLLRLDYN FPHLESLALSLMAG GYQPHLEDFLLRINFN
GCP4_HUMAN	652	
Q642S3 XENLA	652	KYYTOAGGTLGSFGL
GCP4_DROME	654 640	CWFSASHNTSA.
GCP4 ARATH GCP4 MEDTR	724 714	SFYEATARGVLNVVRQRPALPL
A8HPS0_CHLRE Q4WZU3_ASPFU GFH1_SCHPO	725 806	SFYSAKDMDAGGGRASVGVGTAGGRP
GCP2_HUMAN	863	GFYTERLERLSAERSQKATPQVPVLRGPPAPAPRVAVTAQ
GCP3_HUMAN	652	EHINAREFKLRVSLGTRGKRSSHT
GCP5 HUMAN	1021	MEQS
GCP6_HUMAN	1814	NY <mark>Y Q</mark> DA

Supplementary Figure 1 GCPs differ substantially in length and display rather low overall identity/similarity. Structure-based multiple sequence alignment within the GCP4 (upper block) and hGCP (lower block) families. The sequence numbering is for hGCP4. Sequence homology is highlighted in red whereas sequence identity is shown as white letters on a red background. Secondary structure elements (arrows for β -strands and coils for α -helices) of hGCP4 are indicated at the top and colored according to the different subdomains depicted in Figure 1. Residues of GCP4 that are disordered in the crystal structure have been underlined. Percentages of identity/similarity among hGCPs range from 6/11 to 18/36%. Abbreviations: XENLA, *Xenopus laevis*; DANRE, *Danio rerio*; DROME, *Drosophila melanogaster*; ARATH, *Arabidopsis thaliana*; MEDTR, *Medicago truncatula*; CHLRE, *Chlamydomonas reinhardtii*; ASPFU, *Aspergillus fumigatus*; SCHPO, *Schizosaccharomyces pombe*.



Supplementary Figure 2 Structural determinants of the GCP4 fold. (a) Ribbon representation of the overall fold of GCP4 and (b-g) of selected representative interactions found within and between the five different helical bundles and extra secondary structure elements.



Supplementary Figure 3 GCP4 illustrates the structure of the core of all hGCPs. (**a**) Schematic representation of the primary structure of the five hGCPs (light grey). White rectangles indicate deletion regions. Missing loops in the GCP4 structure are in dark grey and corresponding residues are labelled. The most conserved regions are shown with bars color-coded according to the subdomain definition (see Figure 1). The previously defined grip1 and grip2 motifs are boxed in dark grey. (**b**) Ribbon representation in stereo view of the most conserved hGCP regions mapped onto the crystal structure of GCP4.



Supplementary Figure 4 Scatchard plot analysis of GCP4 binding to γ -tubulin. A Kd of 37 ± 23 nM was calculated for this interaction. Two out of four different experiments are represented in this graph. For details, see Materials and Methods.



Supplementary Figure 5 The structure of GCP4 can be used as a template for modeling all GCPs. Ribbon representation of the X-ray structure of GCP4 and of derived models for Spc97 and Spc98, the yeast orthologs of GCPs 2 and 3. Missing loops and insertions of GCP2 and GCP3 with respect to GCP4 are indicated by spheres whose size is proportional to the number of residues assuming a globular fold. Colors are according to the subdomain definition as depicted in Figure 1. The black arrow indicates the location of the 77 and 827 residue-long specific insertions in GCP5 and GCP6, respectively.



Supplementary Figure 6 Interactions between GCP2 and GCP3 within and between γ TuSCs. (a) A molecular model for an entire ring of γ TuSCs, based on the fitting in Figure 4 and the cryo-EM reconstruction of γ TuSC oligomers. (b) Detailed packing of GCP2 and GCP3, both within and between γ TuSCs. Green spheres denote the positions of highly charged inserts in GCP4, which may interfere with lateral assembly of GCP4 in one direction.



Supplementary Figure 7 Normal mode analysis of GCP4 indicates a flex point in the region corresponding to the GCP3 hinge. (a) A side view of the γ TuSC pseudo-atomic model, showing GCP3, with the previously identified hinge region indicated. GCP3 is colored blue in helical bundles 1-3, and cyan in bundles 4 and 5. (b) The first mode from normal mode analysis indicates flexibility at the kink between helical bundles 3 and 4. Structures generated by normal mode analysis are superimposed on the N-terminal part of GCP4.

Ou arrest at must sum	Desidua ranga	Structural neigbours*						
Query structure	Residue failge	PDB†	Z	Rmsd	Lali	Nres	%id	Description
Full range protein	1 654	3ob2	10.1	3.4	114	568/296	8	E3-ubiquitin ligase Cbl (SH2/TBK domain)
Full lange protein	1-034	2jsw	9.5	2.4	106	568/189	9	Talin-1 (Actin-binding domain)
Dundle 1 (D1)	1 50 08 126	1g7d	3.6	2.8	63	79/106	3	ERp29 (C-terminal domain)
Dullule I (DI)	1-30, 98-120	2utg	3.0	3.5	51	79/70	12	Uteroglobin
Dundle 2 (D2)	51 07 127 180	2qsb	5.1	2.8	51	88/85	14	UPF0147 protein
Dullule 2 (D2)	31-97, 127-180	2ox1	4.8	3.7	52	88/62	8	YmgB
Dundle 2 (D2)	191 249	1zx4	4.6	4.6	82	120/180	9	ParB
Dulidle 5 (D5)	181-348	1b47	4.5	3.0	81	120/304	5	Cbl
Pundle 4 (P4)	240 262 461 478 535 572	1k30	5.1	2.8	57	72/363	11	Glycerol-3-phosphate acyltransferase
Dullule 4 (D4)	349-303, 401-478, 333-373	1n5u	4.6	2.2	53	72/583	11	Serum albumin
Pundle 5 (P5)	470 524 574 624	2jsw	10.1	2.1	110	115/189	8	Talin-1 (Actin-binding domain)
Buildle 5 (B5)	479-334, 374-034	3hkk	10.0	2.5	101	115/145	10	Leukotriene C4 synthase
B5 Ctor	170 531 571 651	2jsw	10.1	2.8	112	135/189	10	Talin-1 (Actin-binding domain)
DJ+Clei	479-334, 374-034	3hkk	10.1	2.9	108	135/145	9	Leukotriene C4 synthase
Small damain (SD)	364 460	1ng1	4.4	2.8	66	74/294	5	Ffh (N and GTPase domains)
Siliali dollialii (SD)	304-400	2v3c	4.1	2.9	61	74/403	10	SRP19
B1 B 2	1 180	1zww	5.7	3.4	94	167/196	7	Endophilin-A1 (BAR domain)
D1+D2	1-180	1r0d	5.2	4.6	119	167/190	8	Chimeric IL-8-MGSA protein
$\mathbf{D2} + \mathbf{D2}$	51 07 127 348	2ntz	5.3	4.2	74	208/172	7	ParB
D2+D3	51-97, 127-348	3kph	5.2	3.5	79	208/194	6	Mycoplasma arthritidis-derived mitogen
B3 B 4	181 363 461 478 535 573	1u2m	5.1	4.9	52	192/116	8	Histone-like protein HLP-1
DJ+D4	181–303, 401–478, 333–373	1kw2	4.9	2.6	59	192/455	8	Vitamin D-binding protein
B4+B5+Cter	349 363 461 654	2jsw	10.0	2.8	110	207/189	8	Talin-1 (Actin-binding domain)
	349-303,401-034	1fbv	9.9	3.0	105	207/388	10	c-Cbl
B4+B5+Cter+SD	340 654	1r0d	9.2	3.0	140	281 / 194	8	Chimeric IL-8-MGSA protein
	349-034	2jsw	8.7	3.0	127	281 /189	9	Talin-1 (Actin-binding domain)
Grip 1	163 287	1ybz	4.0	3.8	56	81/76	18	Chorismate mutase
	105-207	2vqe	3.8	2.7	63	81/99	3	30S ribosomal protein S20
Grip 2	337 573	3b60	5.2	6.1	114	214/572	7	MsbA
	337-313	2eho	5.1	4.8	78	214/161	5	GINS complex subunit 4

Supplementary Table 1 Results of Dali search at http://ekhidna.biocenter.helsinki.fi/dali_server/

*For each query structure (not all combinations are shown here), only the first two non-redundant structural neighbours with the highest Z-scores are given. †PDB, PDB identifier of the matched structure; Z, Z-score of the match; Rmsd, rmsd of the match; Lali, number of aligned positions; Nres, number of residues in query/matched structures; %id, sequence identity of aligned positions.

SUPPLEMENTARY DISCUSSION

GCP4 three-dimensional structure

The structure can be viewed as successive layers or bundles (Supplementary Fig. 2a). Starting from the N terminus, the first bundle (residues 1-50, 98-126) comprises helices $\alpha 1$, α 4, and α 5 and part of helix α 2, plus 2 antiparallel β strands (β 1, β 2). This first layer is topped by a second bundle (residues 51-97, 127-180) consisting of the remaining part of helix $\alpha 2$, part of helix $\alpha 8$ and helices $\alpha 3$, $\alpha 6$, and $\alpha 7$, and one disordered loop region spanning 13 residues (Val66-Gln78 between helices $\alpha 2$ and $\alpha 3$). Contiguous to the second bundle is a third bundle (residues 181-348) that includes the second half of helix $\alpha 8$ and helices $\alpha 9$, $\alpha 10$, and $\alpha 11$, plus 2 antiparallel beta strands ($\beta 3$, $\beta 4$). This part of the structure also includes the longest disordered loop (residues Pro209-Pro252 between \$\beta\$ and \$\beta\$) and the partially disordered peptide segment (Gln289-Gly297 between helices α 9 and α 10). The fourth bundle (residues 349-363, 461-478, and 535-573) is made of part of helix α 15 and helices α 12, α 17, and $\alpha 18$ whereas the fifth bundle (residues 479-534, 574-634) encompasses the remaining part of helix $\alpha 15$ and helices $\alpha 16$, $\alpha 19$, and $\alpha 20$. The two C-terminal bundles are flanked on one side by a small domain (residues 364-460) that comprises helices α 13 and α 14 plus 2 antiparallel beta strands (β 5, β 6) and on the other side by the C-terminal helix α 21 (residues 635-654), which runs perpendicular to the diagonal. The small domain also contains one disordered region (Lys423-Pro445 between strands β 5 and β 6). The cohesion of the GCP4 tertiary structure is ensured in part by a set of longitudinal helices interconnecting two adjacent bundles, i.e. helix $\alpha 2$ for bundles 1 and 2, helix $\alpha 8$ for bundles 2 and 3, and helix α 15 for bundles 4 and 5. It is also noteworthy that the major axis of bundle 3 is shifted with respect to the major axes of the first two and the last two bundles, where bundles 1 and 2 on one hand and bundles 3 and 4 on the other hand may be considered a continuation of each other.

The first helix bundle is stabilized by a Leu-rich hydrophobic core that is strongly conserved in the GCP4 family. Peripheral to the hydrophobic core of the first helix bundle, a hydrogen bond involving Tyr100, a highly conserved residue within the GCP4 family, connects the extremities of $\alpha 1$ and $\alpha 4$ (**Supplementary Fig. 2b**). Contiguous to Tyr100, a hydrogen bond between His121 and Glu110 reinforces lateral contacts between helices $\alpha 4$ and $\alpha 5$. The $\beta 1$ - $\beta 2$ anti-parallel β -sheet is docked against helices $\alpha 1$ and $\alpha 2$ through hydrophobic contacts and polar interactions involving the side-chains of Thr50 and Arg54 from helix $\alpha 2$ whereas the hydrophilic side of the sheet is exposed to solvent. A salt bridge

and water-mediated hydrogen bonds between several rather poorly conserved residues (Tyr52, Thr56, Asp94, Gln98, Arg101) pin the extremities of $\alpha 4$ and $\alpha 3$ as well as $\alpha 4$ to $\alpha 2$, thus harnessing together the first two helix bundles. This is reinforced by a set of hydrogen bonds between the side-chains of the non-conserved residues Arg44 (α 2), Asn123 (α 5) and Gln130 ($\alpha 6$). These bonding networks contribute to, or are the result of, the tilts observed between helices $\alpha 3$ and $\alpha 4$ on the one hand and $\alpha 5$ and $\alpha 6$ on the other hand, and the relative orientation between bundles 1 and 2. Unlike the other bundles found in the GCP4 structure, which all contain 4 helices, the second bundle is made of 5 α -helices. It is arranged around a hydrophobic core, enriched in leucine and valine residues conserved in the GCP4 family, where helices $\alpha 3$, $\alpha 6$, $\alpha 7$, and $\alpha 8$ are parallel in pairs while $\alpha 2$ is tilted with respect to the other helices, docked against $\alpha 3$ and $\alpha 6$ that it moves apart (Supplementary Fig. 2a). The topology is under the influence of the tight bonding network that occurs between the top of the second bundle (helices $\alpha 3$, $\alpha 7$, $\alpha 8$ and the loop $\alpha 6$ - $\alpha 7$) and the bottom of the third bundle (α 8 and loops α 8- β 3 and β 4- α 9) (Supplementary Fig. 2c). A hydrogen bond is found between the highly conserved residues Tyr85 and Gln186. The N-terminus of α 3 is capped by a conserved glutamic acid at position 202 within the α 8- β 3 loop. The conformation of this loop is stabilized by the side-chains of two other conserved residues, Asp198 and His149, and a water molecule. Two weakly conserved residues (Gln142 and His160) also form a hydrogen bond, linking $\alpha 6$ and $\alpha 7$ together.

Numerous hydrophobic residues conserved among GCP4 proteins also contribute to the hydrophobic core and to lateral inter-helices interactions found in the third bundle as well as the hydrophobic interface specifically made between helix $\alpha 9$ and the $\beta 3$ - $\beta 4$ anti-parallel β -sheet. Docking of the β -sheet onto the GCP4 structure also pulls away helices $\alpha 8$ and $\alpha 9$. Two salt bridges formed between residues Arg271 and Glu326 and between residues Lys275 and Asp330 pin $\alpha 9$ obliquely to $\alpha 11$ (**Supplementary Fig. 2d**). The strongly conserved His339 tethers the C-terminal side of the $\alpha 9$ - $\alpha 10$ loop to the bundle thereby clamping the extremity of $\alpha 10$ to $\alpha 11$, whereas an additional salt bridge (involving residues Arg313 and Asp324) connects the opposing extremities of $\alpha 10$ and $\alpha 11$ together. A salt bridge formed between the poorly conserved residues His193 and Glu304 anchors $\alpha 8$ to $\alpha 10$. Finally, an aromatic core conserved in GCP4 (Tyr184, Tyr267, Phe320 and Phe325) anchors the bottom of the third bundle and may contribute to the slight curvature of helix $\alpha 8$. The numerous hydrophobic and polar interactions displayed by the third bundle may help preserve the integrity necessary to a putative pivotal role. Interactions between bundle 3 and bundle 4 are

essentially mediated by the C terminus of helix $\alpha 11$ that fits in a shallow hydrophobic groove provided by the N termini of $\alpha 12$ and $\alpha 18$, the short segment connecting $\alpha 17$ to $\alpha 18$, and the 3_{10} helix found at the C-terminal end of the loop $\beta 6 \cdot \alpha 15$ (**Supplementary Fig. 2e**). Polar interactions occur at only two sites on the interface of bundle 3 with the C-terminal domain. The first site involves three relatively conserved residues (Arg333, Asp552, and Glu554) which form salt bridges connecting $\alpha 11$ to $\alpha 18$. On the opposite side, Trp460 forms a hydrogen bond with the main-chain oxygen atom of residue 285, linking the $\beta 6 \cdot \eta 3$ turn to the extremity of $\alpha 9$. Strikingly, the side-chain of Trp460 is at van der Waals contact from Pro461, thus maintaining the aromatic ring of Trp460 in its position. Both Trp460 and Pro461 are almost strictly conserved in the GCP4 family.

The fourth helix bundle comprises four α -helices that establish many hydrophobic interactions through another set of highly conserved aliphatic and aromatic residues arranged around Phe540, Phe563 and the strictly conserved Tyr474. The side chain of Tyr474 forms in turn a hydrogen bond with the conserved His560 within the heart of the hydrophobic core, underlining the importance of these two residues in stabilizing the bundle's fold (**Supplementary Fig. 2e**). In contrast, relatively few electrostatic and polar interactions occur between the helices of the fourth bundle. For example, the only connection between α 12 and α 17 occurs at a salt bridge formed by Lys358 and Glu537, both strongly conserved in the GCP4 family. Gln539, which is conserved polar in all GCP4s, His562, and Asn566 along with three water molecules, form a hydrogen bond network linking α 17 and α 18. Several hydrophobic residues (Val476 and Tyr480 from α 15; Leu568, Phe572, and Phe579 from α 18) delineate an exposed hydrophobic patch at the interface between bundles 4 and 5. This hydrophobic surface could eventually accommodate the long segment between β 5 and β 6 (residues 423-445) and/or the very C-terminal stretch, which are both not seen in the GCP4 structure.

The fifth helix bundle is organized around a hydrophobic core that is capped by Trp495, an almost strictly conserved residue in the GCP4 family, and Phe592 (**Supplementary Fig. 2f**). The bundle is also characterized by an absence of close packing between α 15 and α 16 at the proximal end, and between α 16 and α 20 at the distal end. The gap observed between α 15 and α 16 is nevertheless partially filled by the side-chains of three conserved glutamines (488, 528, and 532), which form hydrogen bonds, and also allows housing helix α 12 (**Supplementary Fig. 2f**). Inter-helical interactions involve the only disulfide bridge of the GCP4 structure, formed between the non-conserved cysteine residues 494 and 589, linking

 α 15 to α 19. A long stretch of conserved residues within the GCP4 structure (526-540), which encompasses the C terminus of α 16, is involved in several polar interactions. A watermediated hydrogen bond network formed among residues Tyr529, Tyr530, Asp534, and Arg633 connects α 16 to α 20. Another hydrogen bond involving His519 and Ser622 also bridges α 16 to α 20 midway along the bundle. Helices α 19 and α 20 both make hydrophobic contacts with α 21. The C-terminal helix is also clamped to the fifth bundle by polar interactions involving the hydroxyl group of the conserved residue Ser631, backbone atoms of the α 20- α 21 loop and the side-chain atom of Asn638, maintaining α 21 perpendicular to α 20.

The two C-terminal bundles also interact with the small domain. This interaction mainly involves helices $\alpha 12$, $\alpha 13$, and $\alpha 15$ and strand $\beta 6$. Fastening of the subdomain onto the fourth bundle is insured by polar interactions involving the side-chains of three conserved residues (Tyr361, Tyr456, and Asn475) as well as that of Gln353 (**Supplementary Fig. 2g**). The interface between the small domain and the fifth helix bundle is smaller. It involves a hydrogen bond formed between the side-chains of two conserved residues, Gln370 and Arg485 (**Supplementary Fig. 2g**). The fold of the subdomain involves a highly conserved aromatic core located on $\alpha 13$ and $\alpha 14$, and onto which strands $\beta 5$ and $\beta 6$ are packed.

A search for structural homology with the program Dali^1 , using the whole GCP4 structure or its domains and any combination as a query structure, revealed that GCP4 represents an original protein fold (Supplementary Table 1). Significant similarity despite very weak sequence homology (less than 10% identity) was obtained between the C-terminal bundle of GCP4 and around 20 proteins containing a similar motif. However, common themes in the use of such α -helical bundle motif cannot be easily foreseen.

The γTuSC model

The location, orientation, and rough boundaries of GCP2 and GCP3 in γ TuSC have been described previously². The GCP4 crystal structure fits very well into both the GCP2 and GCP3 densities (**Fig. 3a**). Indeed, in some regions of the EM map with clear alpha-helical density there is good agreement with the position of GCP4 helices, suggesting a remarkable level of structural conservation (**Fig. 3b**). The kink in the GCP4 structure between helical bundles 3 and 4 closely matches the bent shape of GCP3 in the EM density. The best fit for the straighter GCP2 subunit, however, was achieved by splitting the atomic model in two at the kink point and fitting each half of the structure independently, straightening the structure by about 7°. In the resulting model, an area of empty EM density remains at the base of the

 γ TuSC, which can accommodate the unmodelled N-terminal regions of GCP2 and GCP3, consistent with a prior report that the N-termini cross each other³. The other large area of unfilled EM density, near the second bundle of GCP2, is sufficient to accommodate a yeast-specific GCP2 insertion between helices 6 and 7 (**Fig. 3b**, black arrow). The position of this insert, directly beneath the kink in GCP2, may play a role in the straighter conformation of GCP2 relative to GCP4.

Within the γ TuSC, contacts are made between the first bundles of GCP2 and GCP3, and between their second bundles (**Supplementary Fig. 6**). In the first bundle, α 1, the α 1- β 1 loop, and α 4 of GCP2 pack against α 5 and the N-terminal end of α 2 of GCP3. The second bundles interact through the α 2- α 3 loop of GCP2 and α 6 and the α 7- α 8 loop of GCP3. A more minor contact is made in the third bundle between the α 9- α 10 loop of GCP2 and the Nterminal end of α 11 in GCP3. In γ TuSC oligomers, the equivalent surfaces are used to make lateral contacts between subunits.

- Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. Nucleic Acids Res. 38 Suppl, W545–9 (2010).
- Kollman, J.M., Polka, J.K., Zelter, A., Davis, T.N. & Agard, D.A. Microtubule nucleating gamma-TuSC assembles structures with 13-fold microtubule-like symmetry. *Nature* 466, 879–82 (2010).
- Choy, R.M., Kollman, J.M., Zelter, A., Davis, T.N. & Agard, D.A. Localization and orientation of the gamma-tubulin small complex components using protein tags as labels for single particle EM. J. Struct. Biol. 168, 571–4 (2009).