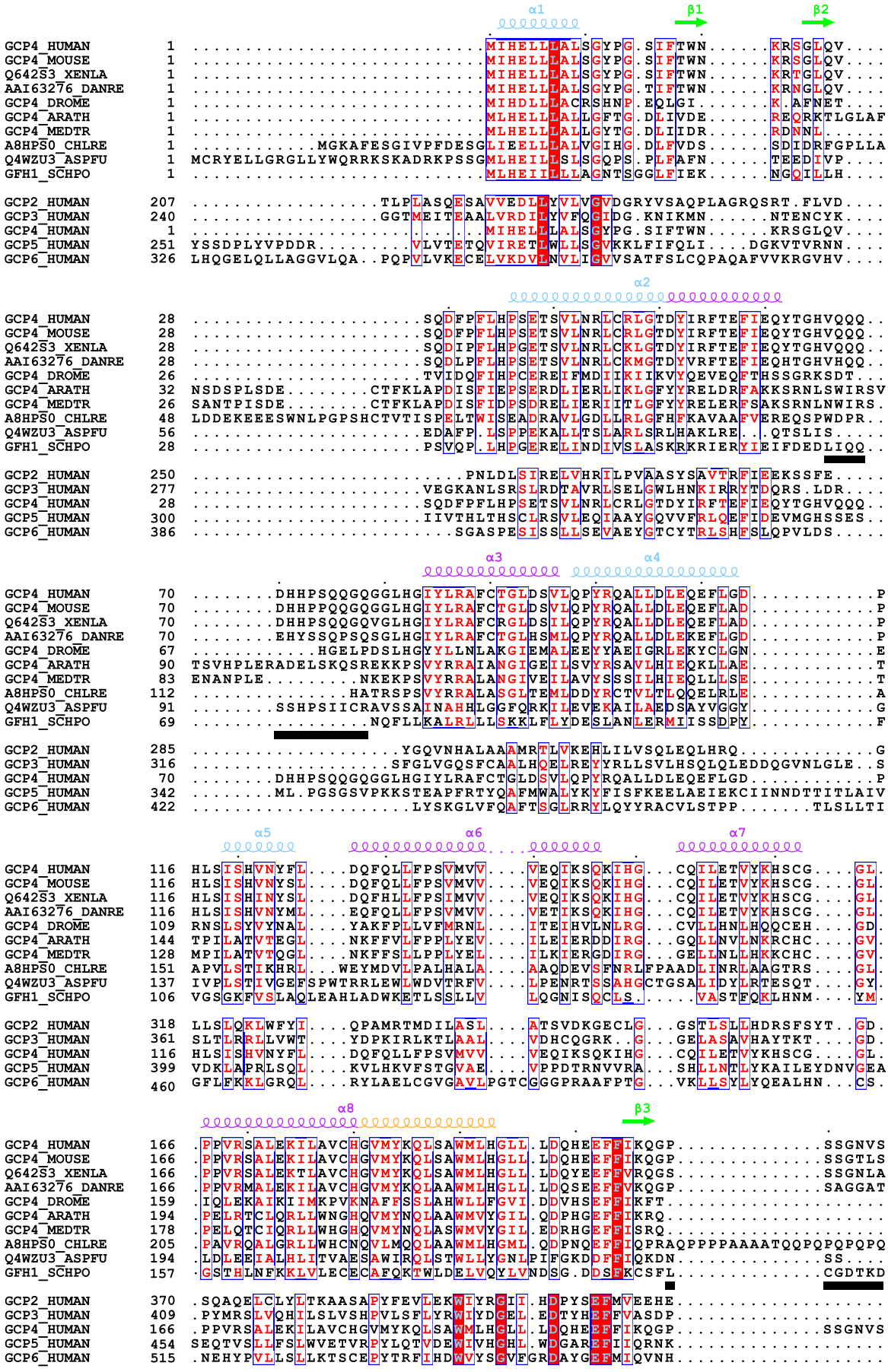
Supplementary information

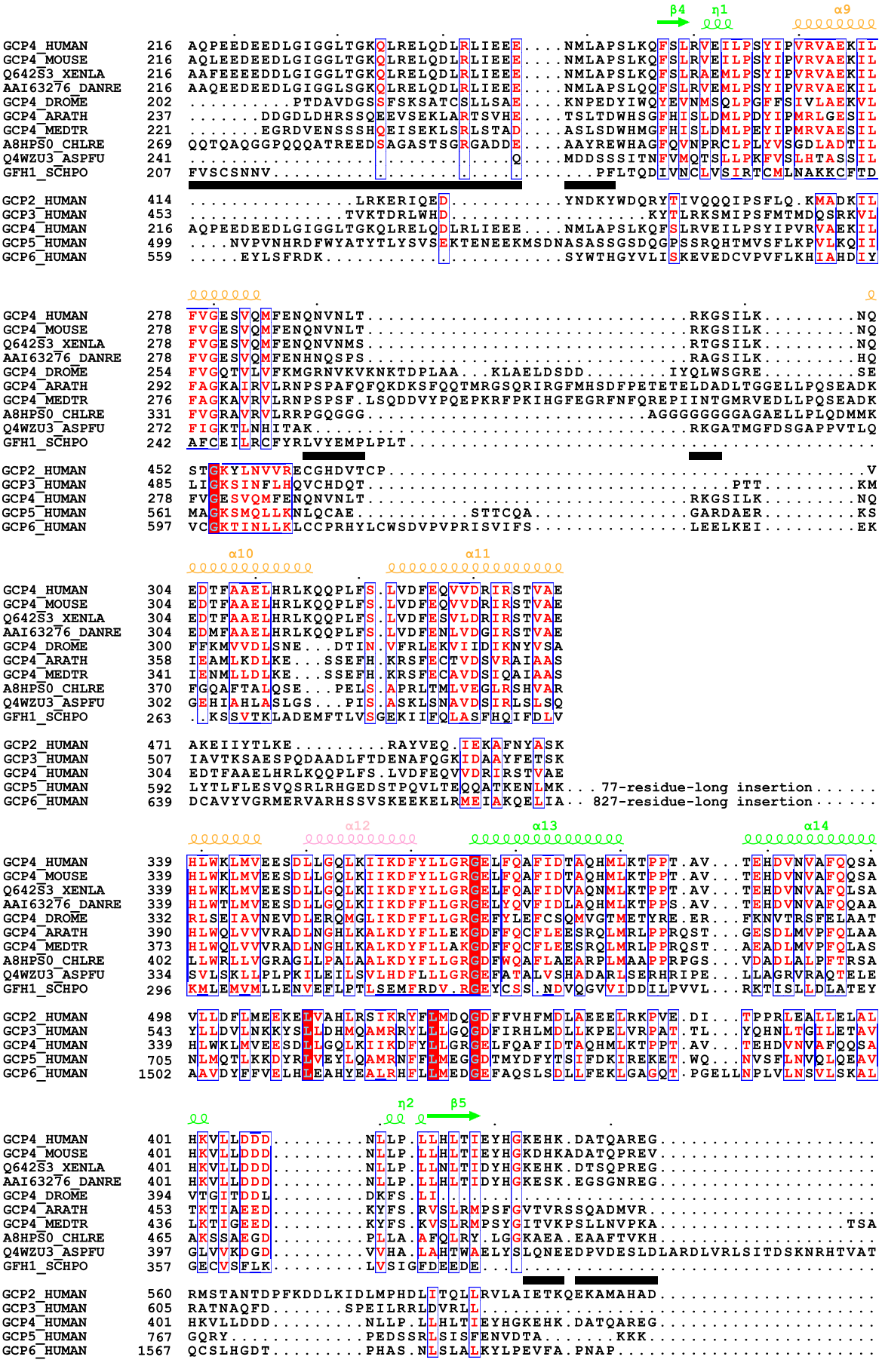
The crystal structure of γ-tubulin complex protein GCP4 provides insight into microtubule nucleation

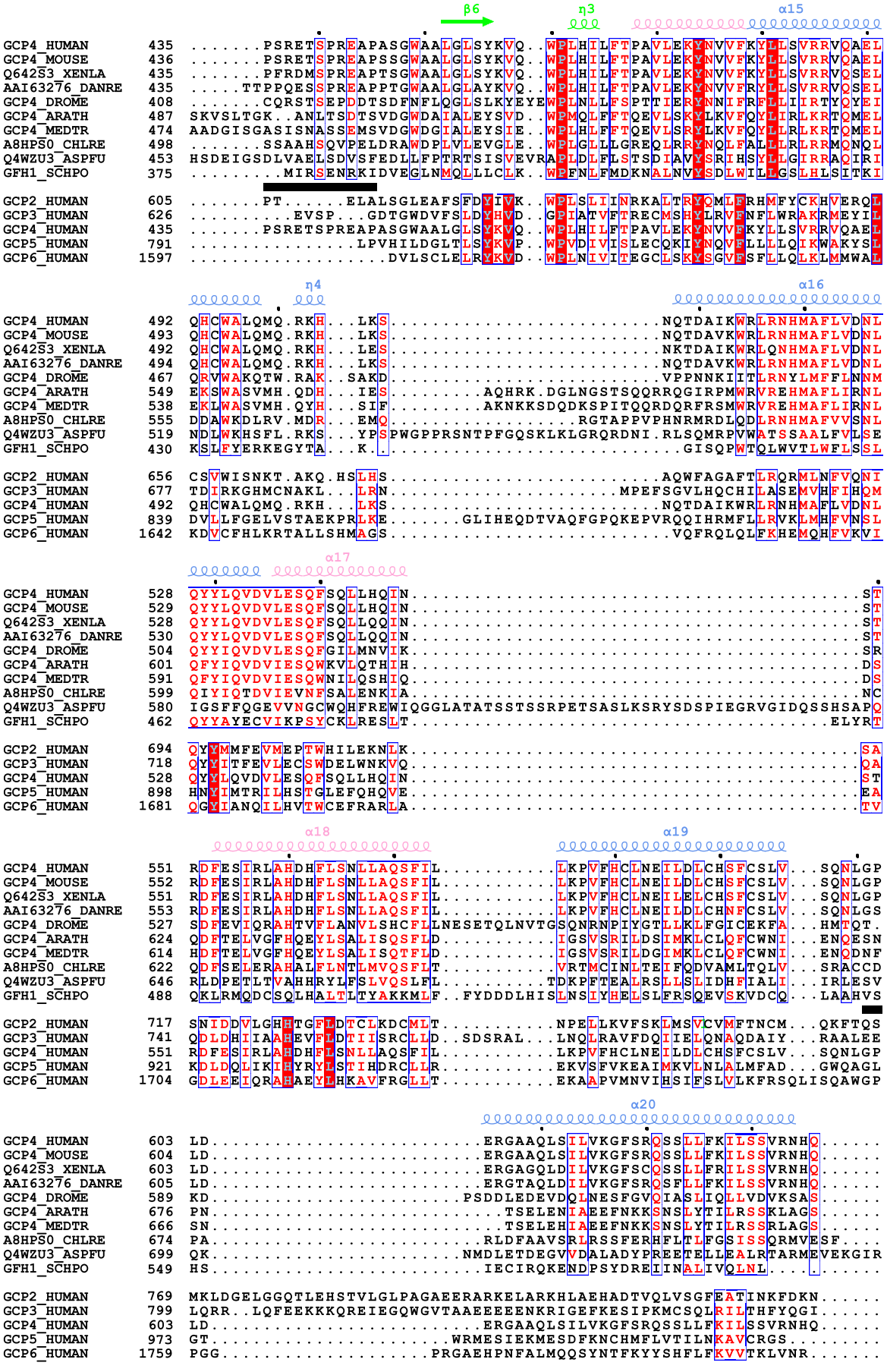
Valérie Guillet[1,2,6], Martine Knibiehler[3,6], Lynn Gregory-Pauron[1,2,6], Marie-Hélène Rémy[3], Cécile Chemin[3], Brigitte Raynaud-Messina[3], Cécile Bon[1,2], Justin M. Kollman[4,5], David A. Agard[4,5], Andreas Merdes[3] & Lionel Mourey[1,2]

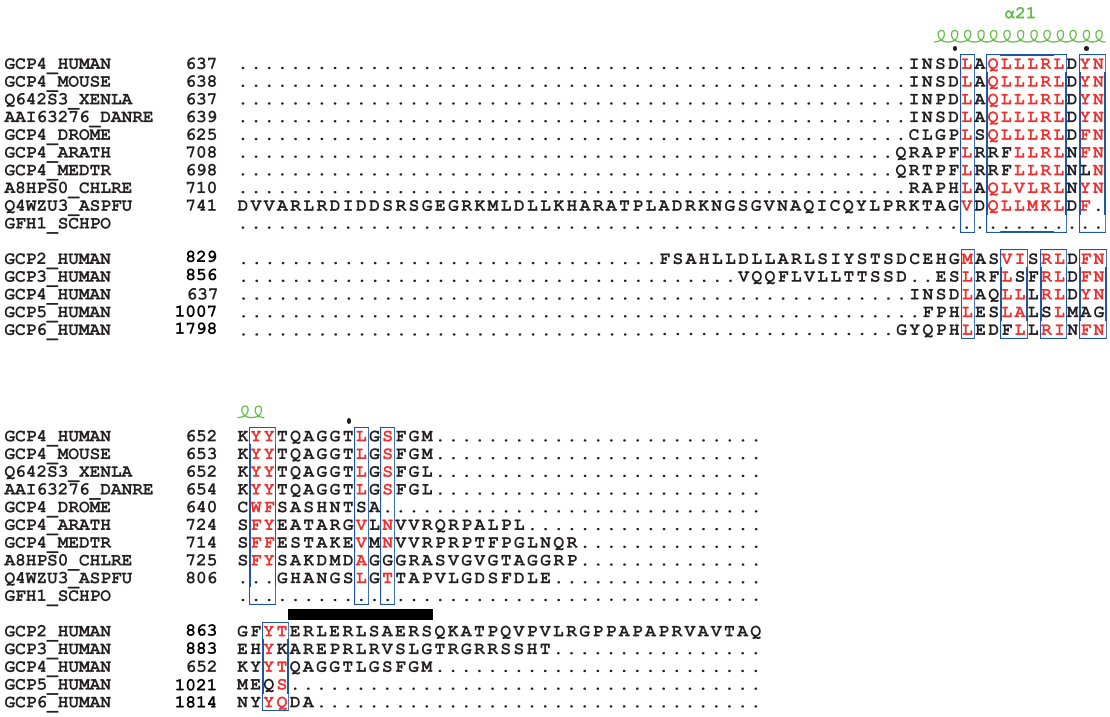
[1]Institut de Pharmacologie et de Biologie Structurale (IPBS), Centre National de la Recherche Scientifique (CNRS), Toulouse, France. [2]Université de Toulouse, Université Paul Sabatier, IPBS, Toulouse, France. [3]Centre de Recherche en Pharmacologie-Santé (CRPS), UMR 2587 CNRS-Pierre Fabre, Toulouse,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



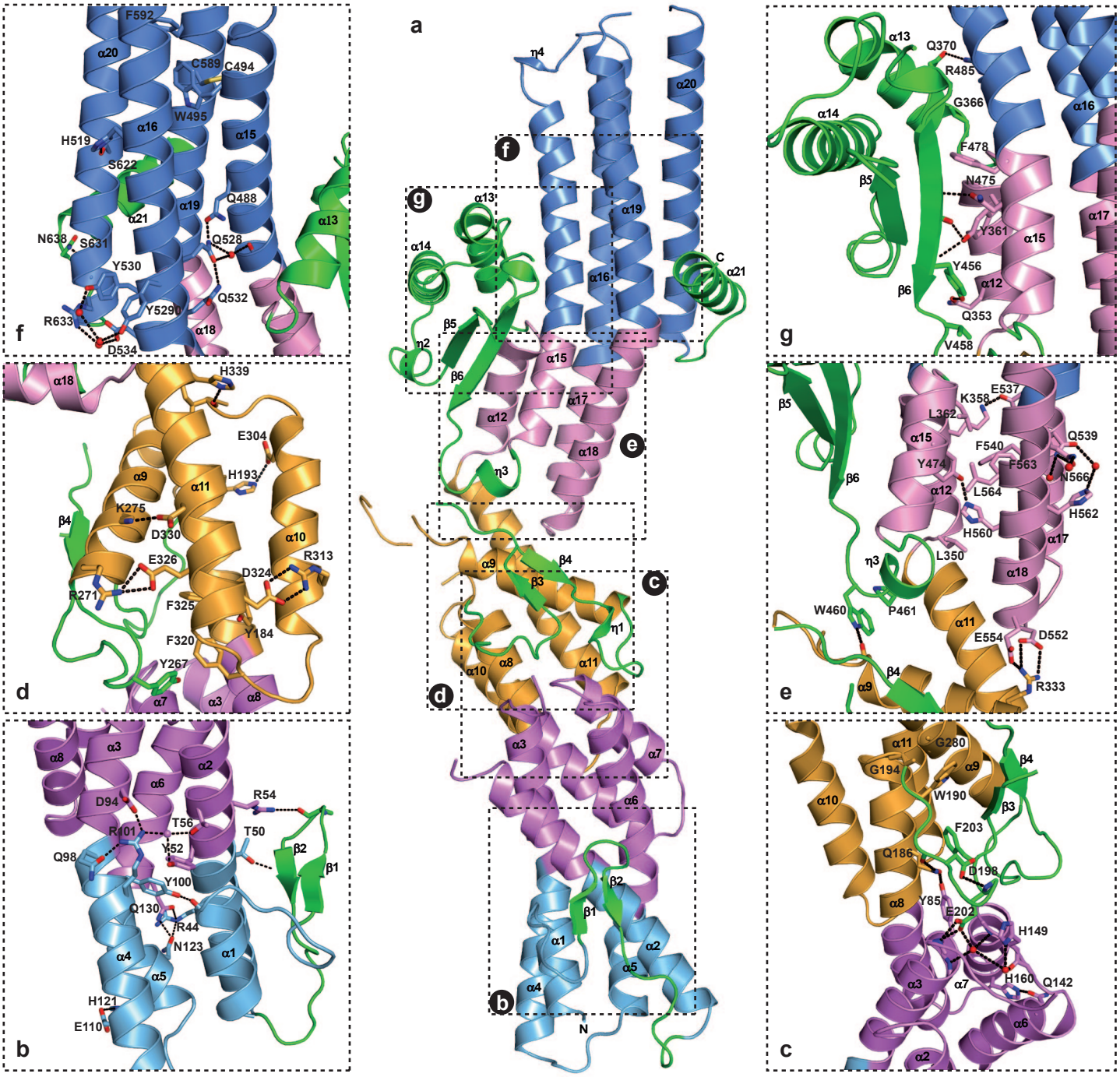
Nature Structural & Molecular Biology: doi:10.1038/nsmb.2083


β4 η1 α9
GCP4_HUMAN 216 AQPEEDEEDLGIGGLTGKQLRELQDLRLIEEE....NMLAPSLKQFSLRVEILPSYIPVRVAEKIL
GCP4_MOUSE 216 AQLEEDEEDLGIGGLTGKQLRELQDLRLIEEE....NMLAPSLKQFSLRVEILPSYIPVRVAEKIL
Q642S3_XENLA 216 AAFEEEEDDLGIGGLTGKQLRELQDLRLIEEE....NMLAPSLKQFSLRAEMLPSYIPVRVAEKIL
AAI63276_DANRE 216 AAQEEDEEDLGIGGLSGKQLRELQDLRLIEEE....NMLAPSLQQFSLRVEMLPSYIPVRVAEKIL
GCP4_DROME 202 .........PTDAVDGSSFSKSATCSLLSAE....KNPEDYIWQYEVNMSQLPGFFSIVLAEKVL
GCP4_ARATH 237 .......DDGDLDHRSSQEEVSEKLARTSVHE....TSLTDWHSGFHISLDMLPDYIPMRLGESIL
GCP4_MEDTR 221 .......EGRDVENSSSHQEISEKLSRLSTAD....ASLSDWHMGFHISLDMLPEYIPMRVAESIL
A8HPS0_CHLRE 269 QQTQAQGGPQQQATREEDSAGASTSGRGADDE....AAYREWHAGFQVNPRCLPLYVSGDLADTIL
Q4WZU3_ASPFU 241 .....Q....MDDSSSITNFVMQTSLLPKFVSLHTASSIL
GFH1_SCHPO 207 FVSCSNNV.....PFLTQDIVNCLVSIRTCMLNAKKCFTD
GCP2_HUMAN 414 .....LRKERIQED.....YNDKYWDQRYTIVQQQIPSFLQ.KMADKIL
GCP3_HUMAN 453 .....TVKTDRLWHD.....KYTLRKSMIPSFMTMDQSRKVL
GCP4_HUMAN 216 AQPEEDEEDLGIGGLTGKQLRELQDLRLIEEE....NMLAPSLKQFSLRVEILPSYIPVRVAEKIL
GCP5_HUMAN 499 ....NVPVNHRDFWYATYTLYSVSEKTENEEKMSDNASASSGSDQGPSSRQHTMVSFLKPVLKQII
GCP6_HUMAN 559 .....EYLSFRDK.....SYWTHGYVLISKEVEDCVPVFLKHIAHDIY
GCP4_HUMAN 278 FVGESVQMFENQNVNLT.....RKGSILK.....NQ
GCP4_MOUSE 278 FVGESVQMFENQNVNLT.....RKGSILK.....NQ
Q642S3_XENLA 278 FVGESVQMFENQNVNMS.....RTGSILK.....NQ
AAI63276_DANRE 278 FVGESVQMFENHNQSPS.....RAGSILK.....HQ
GCP4_DROME 254 FVGQTVLVFKMGRNVKVKNKTDPLAA...KLAELDSDD.....IYQLWSGRE.....SE
GCP4_ARATH 292 FAGKAIRVLRNPSPAFQFQKDKSFQQTMRGSQRIRGFMHSDFPETETELDADLTGGELLPQSEADK
GCP4_MEDTR 276 FAGKAVRVLRNPSPSF.LSQDDVYPQEPKRFPKIHGFEGRFNFQREPIINTGMRVEDLLPQSEADK
A8HPS0_CHLRE 331 FVGRAVRVLRRPGQGGG.....AGGGGGGGGAGAELLPLQDMMK
Q4WZU3_ASPFU 272 FIGKTLNHITAK.....RKGATMGFDSGAPPVTLQ
GFH1_SCHPO 242 AFCEILRCFYRLVYEMPLPLT.....
GCP2_HUMAN 452 STGKYLNVVRECGHDVTCP.....V
GCP3_HUMAN 485 LIGKSINFLHQVCHDQT.....PTT.....KM
GCP4_HUMAN 278 FVGESVQMFENQNVNLT.....RKGSILK.....NQ
GCP5_HUMAN 561 MAGKSMQLLKNLQCAE.....STTCQA.....GARDAER.....KS
GCP6_HUMAN 597 VCGKTINLLKLCCPRHYLCWSDVPVPRISVIFS.....LEELKEI.....EK
α10 α11
GCP4_HUMAN 304 EDTFAAELHRLKQQPLFS.LVDFEQVVDRIRSTVAE
GCP4_MOUSE 304 EDTFAAELHRLKQQPLFS.LVDFEQVVDRIRSTVAE
Q642S3_XENLA 304 EDTFAAELHRLKQQPLFS.LVDFESVLDRIRSTVAE
AAI63276_DANRE 304 EDMFAAELHRLKQQPLFS.LVDFENLVDGIRSTVAE
GCP4_DROME 300 FFKMVVDLSNE...DTIN.VFRLEKVIIDIKNYVSA
GCP4_ARATH 358 IEAMLKDLKE...SSEFH.KRSFECTVDSVRAIAAS
GCP4_MEDTR 341 IENMLLDLKE...SSEFH.KRSFECAVDSIQAIAAS
A8HPS0_CHLRE 370 FGQAFTALQSE...PELS.APRLTMLVEGLRSHVAR
Q4WZU3_ASPFU 302 GEHIAHLASLGS...PIS.ASKLSNAVDSIRLSLSQ
GFH1_SCHPO 263 ..KSSVTKLADEMFTLVSGEKIIFQLASFHQIFDLV
GCP2_HUMAN 471 AKEIIYTLKE.....RAYVEQ.IEKAFNYASK
GCP3_HUMAN 507 IAVTKSAESPQDAADLFTDENAFQGKIDAAYFETSK
GCP4_HUMAN 304 EDTFAAELHRLKQQPLFS.LVDFEQVVDRIRSTVAE
GCP5_HUMAN 592 LYTLFLESVQSRLRHGEDSTPQVLTEQQATKENLMK...77-residue-long insertion.....
GCP6_HUMAN 639 DCAVYVGRMERVARHSSVSKEEKELRMEIAKQELIA..827-residue-long insertion.....
α12 α13 α14
GCP4_HUMAN 339 HLWKLMVEESDLLGQLKIIKDFYLLGRGELFQAFIDTAQHMLKTPPT.AV...TEHDVNVAFQQSA
GCP4_MOUSE 339 HLWKLMVEESDLLGQLKIIKDFYLLGRGELFQAFIDTAQHMLKTPPT.AV...TEHDVNVAFQQSA
Q642S3_XENLA 339 HLWKLMVEESDLLGQLKIIKDFYLLGRGELFQAFIDVAQNMLKTPPT.AV...TEHDVNVAFQLSA
AAI63276_DANRE 339 HLWTLMVEESDLLGQLKIIKDFFLLGRGELYQVFIDLAQHMLKTPPS.AV...TEHDVNVAFQQAA
GCP4_DROME 332 RLSEIAVNEVDLERQMGLIKDFFLLGRGEFYLEFCSQMVGTMETYRE.ER...FKNVTRSFELAAT
GCP4_ARATH 390 HLWQLVVVRADLNGHLKALKDYFLLEKGDFFQCFLEESRQLMRLPPRQST...GESDLMVPFQLAA
GCP4_MEDTR 373 HLWQLVVVRADLNGHLKALKDYFLLAKGDFFQCFLEESRQLMRLPPRQST...AEADLMVPFQLAS
A8HPS0_CHLRE 402 LLWRLLVGRAGLLPALAALKDYFLLGRGDFWQAFLAEARPLMAAPPRPGS...VDADLALPFTRSA
Q4WZU3_ASPFU 334 SVLSKLLPLPKILEILSVLHDFLLLGRGEFATALVSHADARLSERHRIPE...LLAGRVRAQTELE
GFH1_SCHPO 296 KMLEMVMLLENVEFLPTLSEMFRDV.RGEYCSS.NDVQGVVIDDILPVVL...RKTISLLDLATEY
GCP2_HUMAN 498 VLLDFLMEEKEDVAHLRSIKRYFDMDQGGDFFVHFMDLAEEELRKPVE.DI...TPPRLEALLELAL
GCP3_HUMAN 543 YLLDVLNKKYSDLDHMQAMRRYLDLGQGGDFIRHLMDLLKPELVRPAT.TL...YQHNLTGILETAV
GCP4_HUMAN 339 HLWKLMVEESDLLGQLKIIKDFYLLGRGELFQAFIDTAQHMLKTPPT.AV...TEHDVNVAFQQSA
GCP5_HUMAN 705 NLMQTLKKDYRDVEYLQAMRNFFDMEGGDTMYDFYTSIFDKIREKET.WQ...NVSFLNVQLQEAV
GCP6_HUMAN 1502 AAVDYFFVELHDEAHYEALRHFLDMEDGEFAQSLSDLLFEKLGAGQT.PGELLNPLVLNSVLSKAL
η2 β5
GCP4_HUMAN 401 HKVLLDDD.....NLLP.LLHLTIEYHGKEHK.DATQAREG.....
GCP4_MOUSE 401 HKVLLDDD.....NLLP.LLHLTIEYHGKDHKADATQPREV.....
Q642S3_XENLA 401 HKVLLDDD.....NLLP.LLNLTIDYHGKEHK.DTSQPREG.....
AAI63276_DANRE 401 HKVLLDDD.....NLLP.LLHLTIDYHGKESK.EGSGNREG.....
GCP4_DROME 394 VTGITDDL.....DKFS.LI.....
GCP4_ARATH 453 TKTIAEED.....KYFS.RVSLRMPSFGVTVRSSQADMVR.....
GCP4_MEDTR 436 LKTIGEED.....KYFS.KVSLRMPSYGITVKPSLLNVPKA.....TSA
A8HPS0_CHLRE 465 AKSSAEGD.....PLLA.AFQLRY.LGGKAEA.EAAFTVKH.....
Q4WZU3_ASPFU 397 GLVVKDGD.....VVHA.LAHTWAELYSLQNEEDPVDESLDLARDLVRLSITDSKNRHTVAT
GFH1_SCHPO 357 GECVSFLK.....LVSIGFDEEDE.....
GCP2_HUMAN 560 RMSTANTDPFKDDLKIDLMPHDLITQLLRVLAIETKQEKAMAHAD.....
GCP3_HUMAN 605 RATNAQFD.....SPEILRRLDVRLL.....
GCP4_HUMAN 401 HKVLLDDD.....NLLP.LLHLTIEYHGKEHK.DATQAREG.....
GCP5_HUMAN 767 GQRY.....PEDSSRLSISFENVDTA.....KKK.....
GCP6_HUMAN 1567 QCSLHGDT.....PHAS.NLSLALKYLPEVFA.PNAP.....

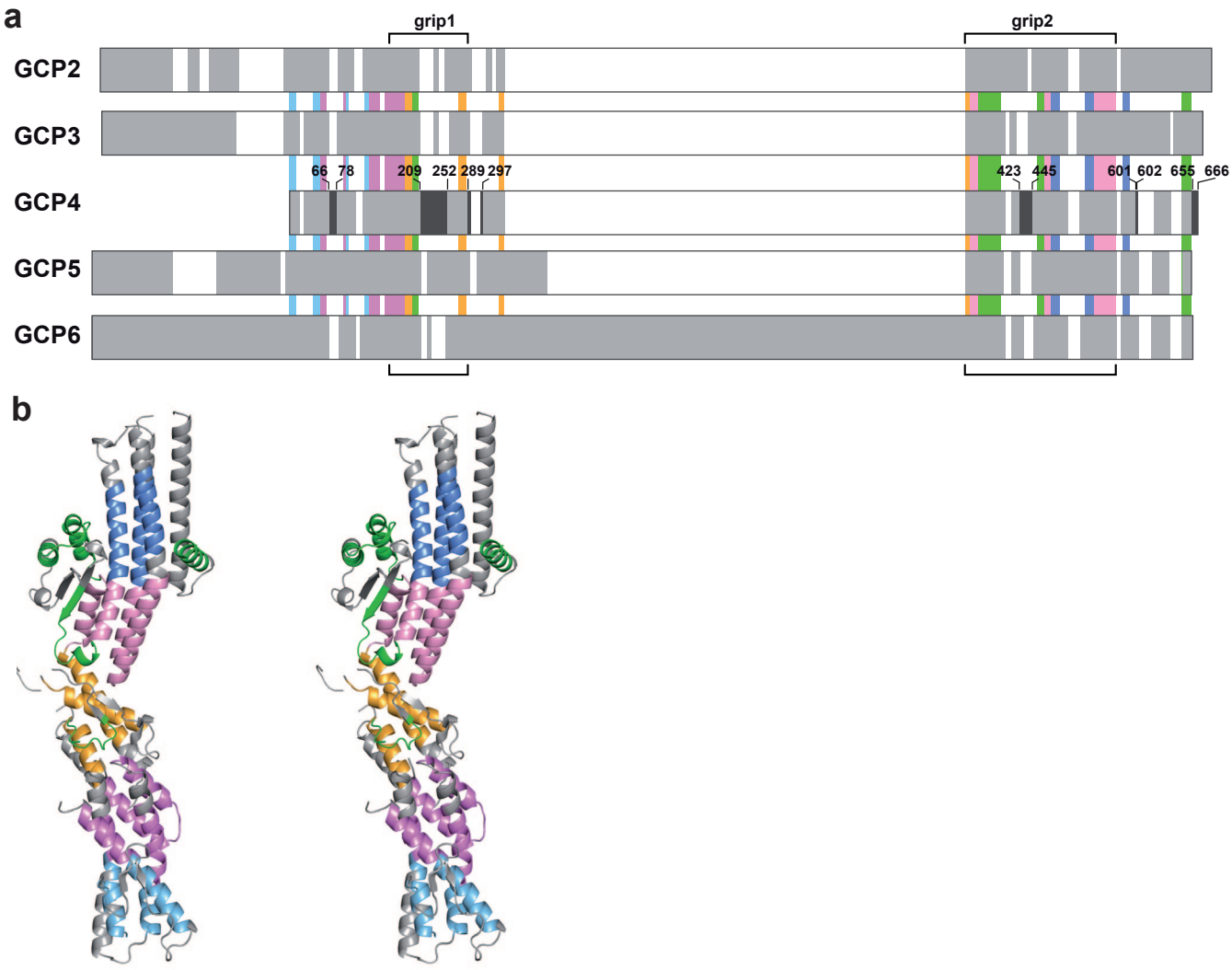

β6
η3
α15
GCP4_HUMAN 435
GCP4_MOUSE 436
Q642S3_XENLA 435
AAI63276_DANRE 435
GCP4_DROME 408
GCP4_ARATH 487
GCP4_MEDTR 474
A8HPS0_CHLRE 498
Q4WZU3_ASPFU 453
GFH1_SCHPO 375
GCP2_HUMAN 605
GCP3_HUMAN 626
GCP4_HUMAN 435
GCP5_HUMAN 791
GCP6_HUMAN 1597
η4
α16
GCP4_HUMAN 492
GCP4_MOUSE 493
Q642S3_XENLA 492
AAI63276_DANRE 494
GCP4_DROME 467
GCP4_ARATH 549
GCP4_MEDTR 538
A8HPS0_CHLRE 555
Q4WZU3_ASPFU 519
GFH1_SCHPO 430
GCP2_HUMAN 656
GCP3_HUMAN 677
GCP4_HUMAN 492
GCP5_HUMAN 839
GCP6_HUMAN 1642
α17
GCP4_HUMAN 528
GCP4_MOUSE 529
Q642S3_XENLA 528
AAI63276_DANRE 530
GCP4_DROME 504
GCP4_ARATH 601
GCP4_MEDTR 591
A8HPS0_CHLRE 599
Q4WZU3_ASPFU 580
GFH1_SCHPO 462
GCP2_HUMAN 694
GCP3_HUMAN 718
GCP4_HUMAN 528
GCP5_HUMAN 898
GCP6_HUMAN 1681
α18
α19
GCP4_HUMAN 551
GCP4_MOUSE 552
Q642S3_XENLA 551
AAI63276_DANRE 553
GCP4_DROME 527
GCP4_ARATH 624
GCP4_MEDTR 614
A8HPS0_CHLRE 622
Q4WZU3_ASPFU 646
GFH1_SCHPO 488
GCP2_HUMAN 717
GCP3_HUMAN 741
GCP4_HUMAN 551
GCP5_HUMAN 921
GCP6_HUMAN 1704
α20
GCP4_HUMAN 603
GCP4_MOUSE 604
Q642S3_XENLA 603
AAI63276_DANRE 605
GCP4_DROME 589
GCP4_ARATH 676
GCP4_MEDTR 666
A8HPS0_CHLRE 674
Q4WZU3_ASPFU 699
GFH1_SCHPO 549
GCP2_HUMAN 769
GCP3_HUMAN 799
GCP4_HUMAN 603
GCP5_HUMAN 973
GCP6_HUMAN 1759



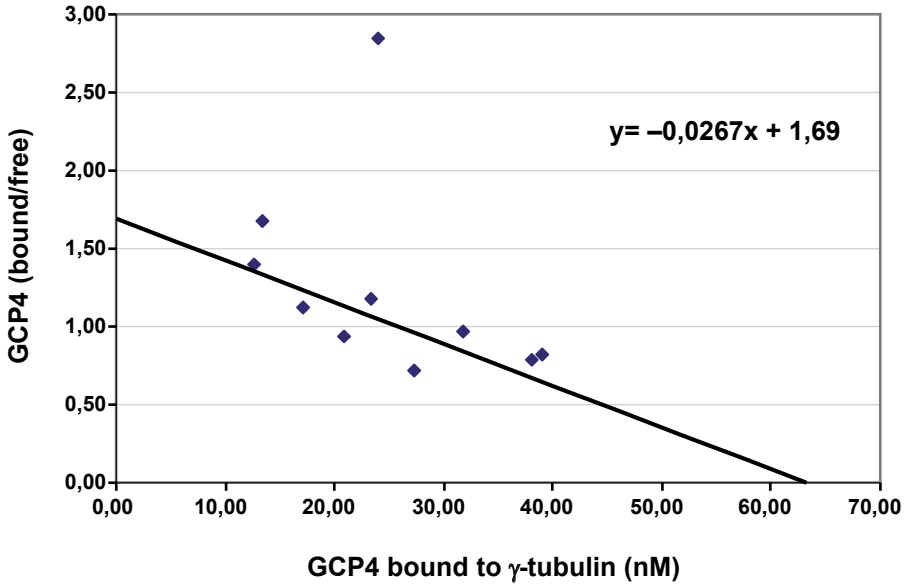
Supplementary Figure 1 GCPs differ substantially in length and display rather low overall identity/similarity. Structure-based multiple sequence alignment within the GCP4 (upper block) and hGCP (lower block) families. The sequence numbering is for hGCP4. Sequence homology is highlighted in red whereas 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*.



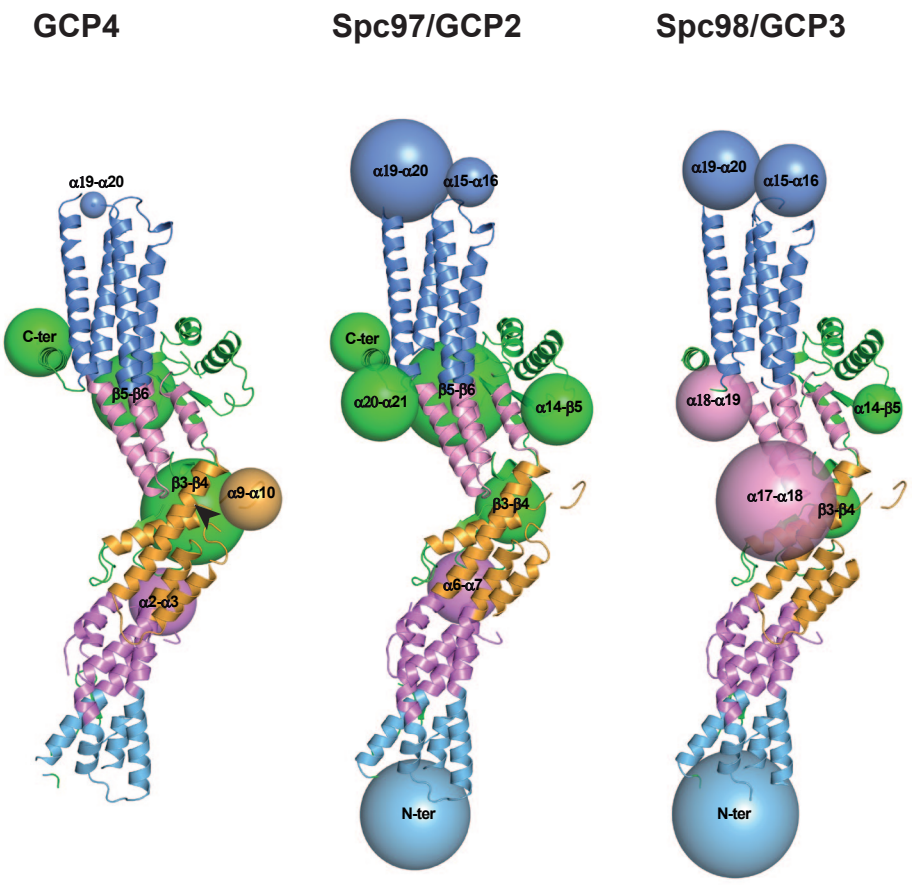
Supplementary Figure 2 Structural determinants of the GCP4 fold. (**a**) Ribbon representation of the overall fold of GCP4 and (**b-g**) of selected representative interactions found within and between the five different helical bundles and extra secondary structure elements.



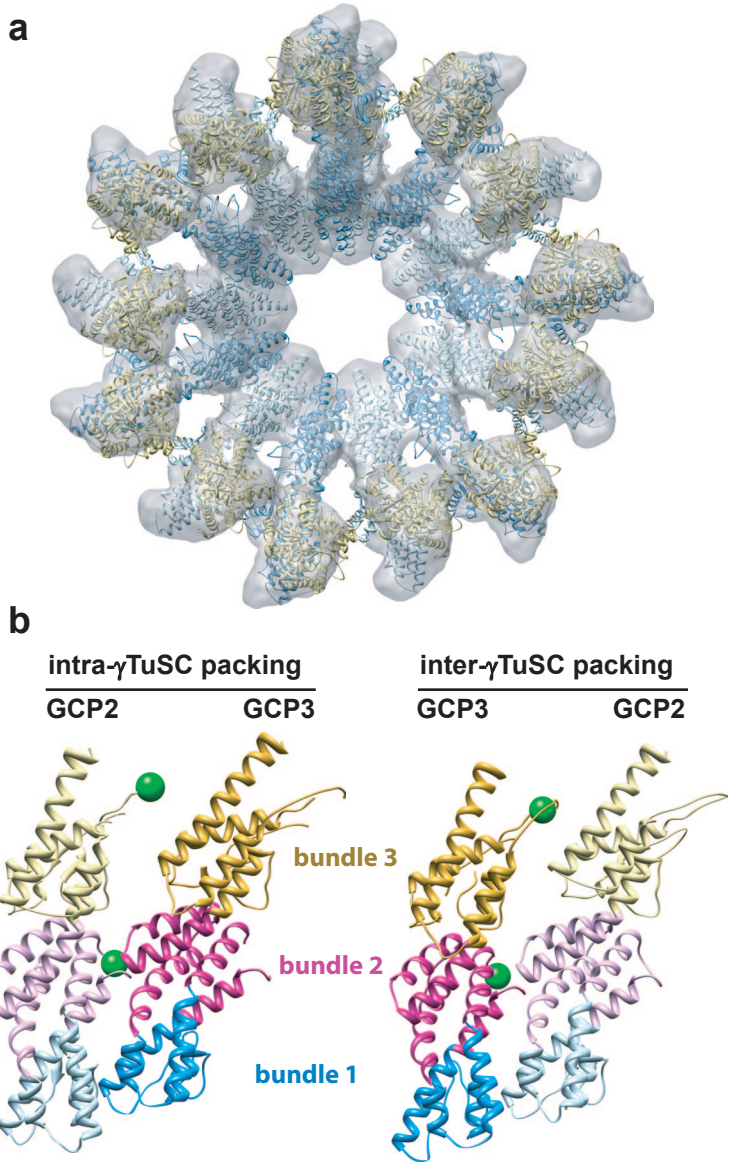
Supplementary Figure 3 GCP4 illustrates the structure of the core of all hGCPs. (**a**) Schematic representation of the primary structure of the five hGCPs (light grey). White rectangles indicate deletion regions. Missing loops in the GCP4 structure are in dark grey and corresponding residues are labelled. 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.



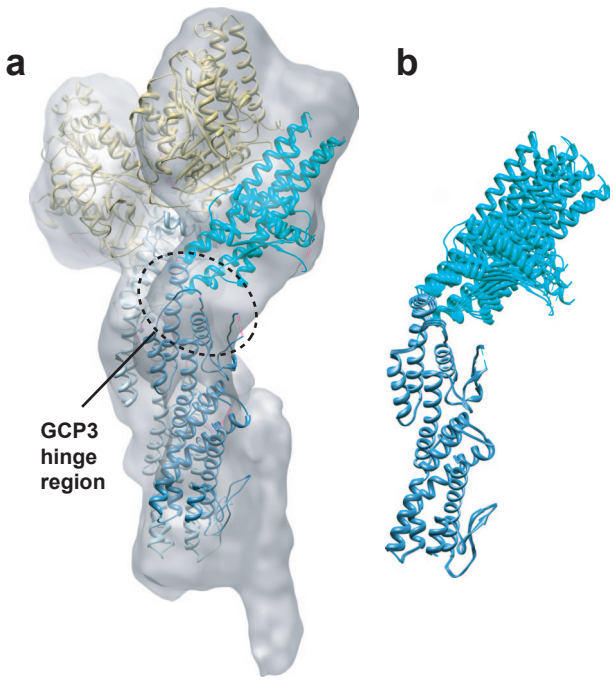
Supplementary Figure 4 Scatchard plot analysis of GCP4 binding to γ-tubulin. A Kd of 37 ± 23 nM was calculated for this interaction. Two out of four different experiments are represented in this graph. For details, see Materials and Methods.



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 γTuSCs. (**a**) A molecular model for an entire ring of γTuSCs, based on the fitting in Figure 4 and the cryo-EM reconstruction of γTuSC oligomers. (**b**) Detailed packing of GCP2 and GCP3, both within and between γTuSCs. 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 γTuSC 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.

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

Query structure	Residue range	Structural neigbours*						
		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)**
		2jsw	**9.5**	**2.4**	**106**	**568/189**	**9**	**Talin-1 (Actin-binding domain)**
Bundle 1 (B1)	1–50, 98–126	1g7d	3.6	2.8	63	79/106	3	ERp29 (C-terminal domain)
		2utg	3.0	3.5	51	79/70	12	Uteroglobin
Bundle 2 (B2)	51–97, 127–180	2qsb	5.1	2.8	51	88/85	14	UPF0147 protein
		2oxl	4.8	3.7	52	88/62	8	YmgB
Bundle 3 (B3)	181–348	1zx4	4.6	4.6	82	120/180	9	ParB
		1b47	4.5	3.0	81	120/304	5	Cbl
Bundle 4 (B4)	349–363, 461–478, 535–573	1k30	5.1	2.8	57	72/363	11	Glycerol-3-phosphate acyltransferase
		1n5u	4.6	2.2	53	72/583	11	Serum albumin
Bundle 5 (B5)	479–534, 574–634	**2jsw**	**10.1**	**2.1**	**110**	**115/189**	**8**	**Talin-1 (Actin-binding domain)**
		3hkk	**10.0**	**2.5**	**101**	**115/145**	**10**	**Leukotriene C4 synthase**
B5+Cter	479–534, 574–654	**2jsw**	**10.1**	**2.8**	**112**	**135/189**	**10**	**Talin-1 (Actin-binding domain)**
		3hkk	**10.1**	**2.9**	**108**	**135/145**	**9**	**Leukotriene C4 synthase**
Small domain (SD)	364–460	1ng1	4.4	2.8	66	74/294	5	Ffh (N and GTPase domains)
		2v3c	4.1	2.9	61	74/403	10	SRP19
B1+B2	1–180	1zww	5.7	3.4	94	167/196	7	Endophilin-A1 (BAR domain)
		1r0d	5.2	4.6	119	167/190	8	Chimeric IL-8-MGSA protein
B2+B3	51–97, 127–348	2ntz	5.3	4.2	74	208/172	7	ParB
		3kph	5.2	3.5	79	208/194	6	Mycoplasma arthritidis-derived mitogen
B3+B4	181–363, 461–478, 535–573	1u2m	5.1	4.9	52	192/116	8	Histone-like protein HLP-1
		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)**
		1fbv	**9.9**	**3.0**	**105**	**207/388**	**10**	**c-Cbl**
B4+B5+Cter+SD	349–654	**1r0d**	**9.2**	**3.0**	**140**	**281 /194**	**8**	**Chimeric IL-8-MGSA protein**
		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
		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
		2eho	5.1	4.8	78	214/161	5	GINS complex subunit 4

*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 α1, α4, and α5 and part of helix α2, plus 2 antiparallel β strands (β1, β2). This first layer is topped by a second bundle (residues 51-97, 127-180) consisting of the remaining part of helix α2, part of helix α8 and helices α3, α6, and α7, and one disordered loop region spanning 13 residues (Val66-Gln78 between helices α2 and α3). Contiguous to the second bundle is a third bundle (residues 181-348) that includes the second half of helix α8 and helices α9, α10, and α11, plus 2 antiparallel beta strands (β3, β4). This part of the structure also includes the longest disordered loop (residues Pro209-Pro252 between β3 and β4) and the partially disordered peptide segment (Gln289-Gly297 between helices α9 and α10). The fourth bundle (residues 349-363, 461-478, and 535-573) is made of part of helix α15 and helices α12, α17, and α18 whereas the fifth bundle (residues 479-534, 574-634) encompasses the remaining part of helix α15 and helices α16, α19, and α20. The two C-terminal bundles are flanked on one side by a small domain (residues 364-460) that comprises helices α13 and α14 plus 2 antiparallel beta strands (β5, β6) and on the other side by the C-terminal helix α21 (residues 635-654), which runs perpendicular to the diagonal. The small domain also contains one disordered region (Lys423-Pro445 between strands β5 and β6). The cohesion of the GCP4 tertiary structure is ensured in part by a set of longitudinal helices interconnecting two adjacent bundles, i.e. helix α2 for bundles 1 and 2, helix α8 for bundles 2 and 3, and helix α15 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 α1 and α4 (**Supplementary Fig. 2b**). Contiguous to Tyr100, a hydrogen bond between His121 and Glu110 reinforces lateral contacts between helices α4 and α5. The β1-β2 anti-parallel β-sheet is docked against helices α1 and α2 through hydrophobic contacts and polar interactions involving the side-chains of Thr50 and Arg54 from helix α2 whereas the hydrophilic side of the sheet is exposed to solvent. A salt bridge

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

and water-mediated hydrogen bonds between several rather poorly conserved residues (Tyr52, Thr56, Asp94, Gln98, Arg101) pin the extremities of α4 and α3 as well as α4 to α2, 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 (α2), Asn123 (α5) and Gln130 (α6). These bonding networks contribute to, or are the result of, the tilts observed between helices α3 and α4 on the one hand and α5 and α6 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 α3, α6, α7, and α8 are parallel in pairs while α2 is tilted with respect to the other helices, docked against α3 and α6 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 α3, α7, α8 and the loop α6-α7) and the bottom of the third bundle (α8 and loops α8-β3 and β4-α9) (**Supplementary Fig. 2c**). A hydrogen bond is found between the highly conserved residues Tyr85 and Gln186. The N-terminus of α3 is capped by a conserved glutamic acid at position 202 within the α8-β3 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 α6 and α7 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 α9 and the β3-β4 anti-parallel β-sheet. Docking of the β-sheet onto the GCP4 structure also pulls away helices α8 and α9. Two salt bridges formed between residues Arg271 and Glu326 and between residues Lys275 and Asp330 pin α9 obliquely to α11 (**Supplementary Fig. 2d**). The strongly conserved His339 tethers the C-terminal side of the α9-α10 loop to the bundle thereby clamping the extremity of α10 to α11, whereas an additional salt bridge (involving residues Arg313 and Asp324) connects the opposing extremities of α10 and α11 together. A salt bridge formed between the poorly conserved residues His193 and Glu304 anchors α8 to α10. 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 α8. 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

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

essentially mediated by the C terminus of helix α11 that fits in a shallow hydrophobic groove provided by the N termini of α12 and α18, the short segment connecting α17 to α18, and the 3_{10} helix found at the C-terminal end of the loop β6-α15 (**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 α11 to α18. On the opposite side, Trp460 forms a hydrogen bond with the main-chain oxygen atom of residue 285, linking the β6-η3 turn to the extremity of α9. 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 α12 and α17 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 α17 and α18. Several hydrophobic residues (Val476 and Tyr480 from α15; Leu568, Phe572, and Phe579 from α18) delineate an exposed hydrophobic patch at the interface between bundles 4 and 5. This hydrophobic surface could eventually accommodate the long segment between β5 and β6 (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 α15 and α16 at the proximal end, and between α16 and α20 at the distal end. The gap observed between α15 and α16 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 α12 (**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

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

α15 to α19. A long stretch of conserved residues within the GCP4 structure (526-540), which encompasses the C terminus of α16, is involved in several polar interactions. A water-mediated hydrogen bond network formed among residues Tyr529, Tyr530, Asp534, and Arg633 connects α16 to α20. Another hydrogen bond involving His519 and Ser622 also bridges α16 to α20 midway along the bundle. Helices α19 and α20 both make hydrophobic contacts with α21. 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 α20-α21 loop and the side-chain atom of Asn638, maintaining α21 perpendicular to α20.

The two C-terminal bundles also interact with the small domain. This interaction mainly involves helices α12, α13, and α15 and strand β6. 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 α13 and α14, and onto which strands β5 and β6 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 γTuSC have been described previously[2]. 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

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

γTuSC, which can accommodate the unmodelled N-terminal regions of GCP2 and GCP3, consistent with a prior report that the N-termini cross each other[3]. 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 γTuSC, contacts are made between the first bundles of GCP2 and GCP3, and between their second bundles (**Supplementary Fig. 6**). In the first bundle, α1, the α1-β1 loop, and α4 of GCP2 pack against α5 and the N-terminal end of α2 of GCP3. The second bundles interact through the α2-α3 loop of GCP2 and α6 and the α7-α8 loop of GCP3. A more minor contact is made in the third bundle between the α9-α10 loop of GCP2 and the N-terminal end of α11 in GCP3. In γTuSC oligomers, the equivalent surfaces are used to make lateral contacts between subunits.

1. Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38** Suppl, W545–9 (2010).
2. 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).
3. 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).

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