SCIENTIFIC REP⚙RTS

OPEN

ACCORD: an assessment tool to determine the orientation of homodimeric coiled-coils

Byeong-Won Kim[1], Yang Ouk Jung[1], Min Kyung Kim[1], Do Hoon Kwon[1], Si Hoon Park[1], Jun Hoe Kim[1], Yong-Boo Kuk[1], Sun-Joo Oh[1], Leehyeon Kim[1], Bong Heon Kim[1], Woo Seok Yang[1] & Hyun Kyu Song[1,2]

The coiled-coil (CC) domain is a very important structural unit of proteins that plays critical roles in various biological functions. The major oligomeric state of CCs is a dimer, which can be either parallel or antiparallel. The orientation of each α-helix in a CC domain is critical for the molecular function of CC-containing proteins, but cannot be determined easily by sequence-based prediction. We developed a biochemical method for assessing differences between parallel and antiparallel CC homodimers and named it ACCORD (**A**ssessment tool for homodimeric **C**oiled-**C**oil **OR**ientation **D**ecision). To validate this technique, we applied it to 15 different CC proteins with known structures, and the ACCORD results identified these proteins well, especially with long CCs. Furthermore, ACCORD was able to accurately determine the orientation of a CC domain of unknown directionality that was subsequently confirmed by X-ray crystallography and small angle X-ray scattering. Thus, ACCORD can be used as a tool to determine CC directionality to supplement the results of *in silico* prediction.

A coiled-coil domain (CCD) consists of two or more α-helices that are twisted around each other to form a superhelical structure[1]. Coiled-coils (CCs) have a heptad repeat pattern (a-b-c-d-e-f-g)$_n$, in which the *a* and *d* positions are usually occupied by hydrophobic residues. Based on this sequence feature, CCs are easily detected by prediction programs or servers such as COILS[2], Marcoil[3], Paircoil2 (ref. 4), and SOCKET[5]. However, they are very versatile structural domains that can adopt many different structures[1]. They can form different oligomeric complexes, and the helix orientation can either be parallel or antiparallel. Because of these characteristics, CCs participate in diverse protein-protein interactions that are involved in many biological functions and components such as fibrous filament formation[6], gene regulation[7,8], development of cartilage and bone[9], membrane channels[10], protein degradation[11,12], protein trafficking[13–15], and molecular chaperones[16]. To understand the detailed molecular functions of CC-containing proteins, it is absolutely necessary to determine the relative orientation of each coil, because the functional domain outside of a parallel CCD dimer must be in close proximity, whereas that outside an antiparallel CCD is located further away. Many programs and servers have been developed to determine the oligomeric state of CCDs; these include SCORER 2.0 (ref. 17), PrOCoil[18], Multicoil2 (ref. 19), and LOGICOIL[20]. However, none of these programs adequately discriminates between the parallel and antiparallel orientations.

As an alternative to *in silico* prediction of CC orientation, the introduction of a cysteine residue at an appropriate site in a helix can be used to determine the CC orientation by analysing the resulting intermolecular disulphide bridge[21,22]. However, although strategies to rapidly evaluate sequence-stability relationships in the parallel coiled-coil motif have been reported[23,24], this approach has the following limitations: first, a cysteine mutation in a CCD often cannot form a proper disulphide bond because of incorrect positioning, overall structural perturbation, and/or oxidation of the sulphur atom[25–27]; and second, the CC molecule itself is sometimes not properly assembled in a heterologous expression system[28] and shows low solubility (Supplementary Fig. 1).

To overcome the above limitations, we developed a new biochemical technique using a fusion tag, which is an appropriately spaced and oriented molecule that enhances the solubility of CC proteins, to assess the orientation of the CC dimer.

[1]Division of Life Sciences, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Korea. [2]Center for Molecular Dynamics and Spectroscopy, Institute of Basic Science, Seoul 02841, Korea. Correspondence and requests for materials should be addressed to H.K.S. (email: hksong@korea.ac.kr)
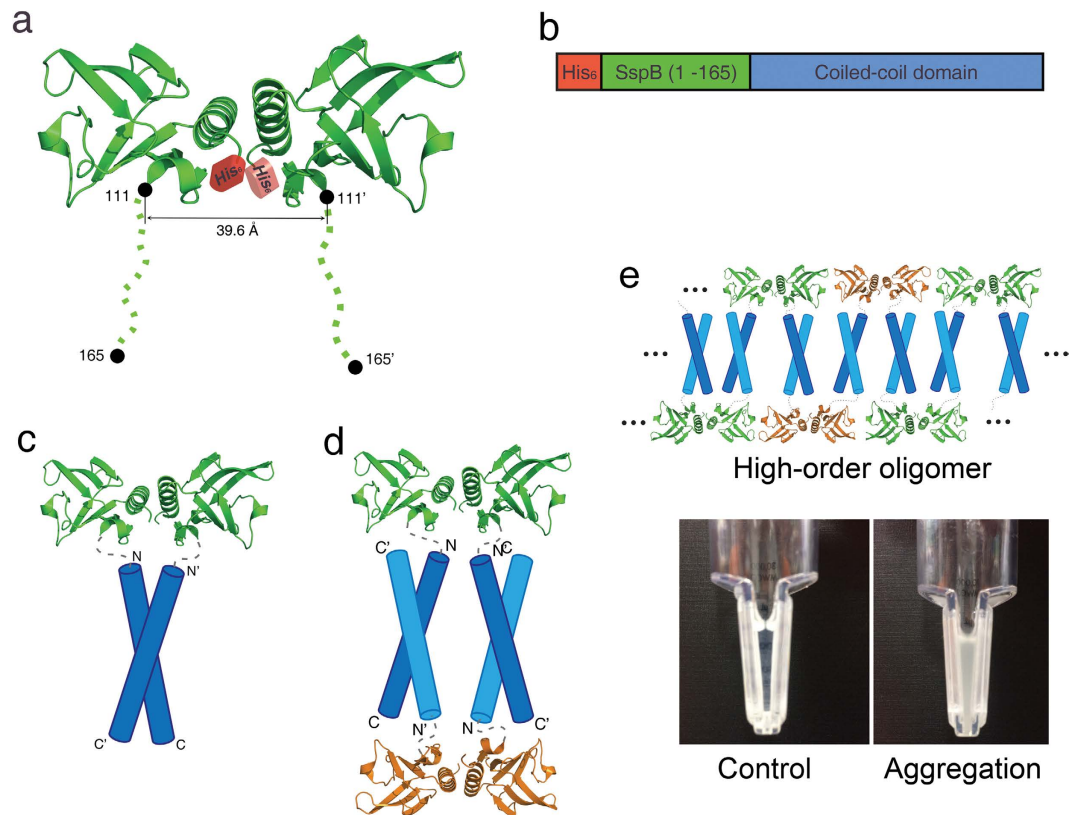
Figure 1. Design of the stringent starvation protein B (SspB)-fusion protein. (**a**) Overall structure of the SspB protein. The tail residues from 112 to 165, shown in dots, were not modelled because of their flexibility. The distance between Asp111 and Asp111′ is approximately 40 Å. The N-terminal hexa-histidine tag is shown as a red block. (**b**) Domain architecture of the SspB-CCD fusion protein. The N-terminal hexa-histidine tag, full-length SspB (residues 1 to 165), and target coiled-coil domain (CCD) are coloured red, green, and blue, respectively. (**c**) Schematic model of a dimer of SspB fused with a parallel CCD. The N- and C-termini of the CCs are indicated. (**d**) Schematic model of a possible arrangement of a tetramer of SspB fused with an antiparallel CCD. The N- and C-termini of the CCs are indicated. (**e**) A possible arrangement of higher-order oligomers of SspB fused with an antiparallel CCD and protein aggregation during concentration of the sample (left: soluble control protein; right: aggregation of higher-order oligomers).

Results

Stringent starvation protein B (SspB) is an appropriate fusion tag. As a fusion tag, we used SspB, which is an adaptor protein that delivers ssrA-tagged proteins to the ClpXP degradation machine in *Escherichia coli*[29,30]. SspB is a dimeric protein with two long and flexible C-terminal tails, both of which extend in the same direction (Fig. 1a)[31–33]. For the purpose of purification, hexa-histidine residues were added to the N-terminus, and the target CCD was attached to the C-terminus of full-length SspB (residues 1–165, with a 55-amino acid residue long C-terminal tail) (Fig. 1b). The C-terminal tail has enough degrees of freedom to accommodate each coil in the CCD. The idea is simple and is as follows: If the CC dimer is oriented in a parallel manner, the SspB-CC fusion protein behaves as a dimer in solution, as shown in Fig. 1c; however, if the CC dimer is oriented in an antiparallel manner, the SspB-CC fusion protein behaves as a soluble tetramer (Fig. 1d) or forms higher-order oligomers and/or aggregates (Fig. 1e).

Size-exclusion chromatography with multi-angle light scattering (SEC-MALS) analyses of parallel or antiparallel CC proteins. To validate this approach, we selected several parallel and antiparallel CC proteins whose structures have already been solved. The length of the CCDs varied from relatively short (49.7 Å) to long (225.7 Å; Figs 2a–h and 3a–f). In total, 14 parallel and antiparallel CC proteins were cloned into expression vectors, and the expressed fusion proteins were analysed by SEC-MALS. Molecular weight (MW) determination by MALS is critical, since dimeric CC proteins usually elute earlier than expected in SEC because of their elongated shapes. Intriguingly, the results were significantly different between parallel and antiparallel CCDs. Parallel CC proteins (GCN4, APC, Atg16, SCOC, LRRFIP1, Ndel1, TPM1, and ROCK1; see the legend in Table 1 for the full names of these proteins) were clearly dimeric in solution (Fig. 4a–h). When fused with SspB, the CCD of ROCK1 kinase, which is very long (225.7 Å)[34], behaves as a 95-kDa dimer in solution (Fig. 4h,i). It must be noted that the distance between the two N-termini of parallel CCDs was in the range of 5.8–23.6 Å, which might fit well within the flexible C-terminal tails of SspB. In rare cases, the presence of a different oligomeric state was detected, but the major peak always corresponded to a dimer of parallel CCDs (Fig. 4f,g). This is an
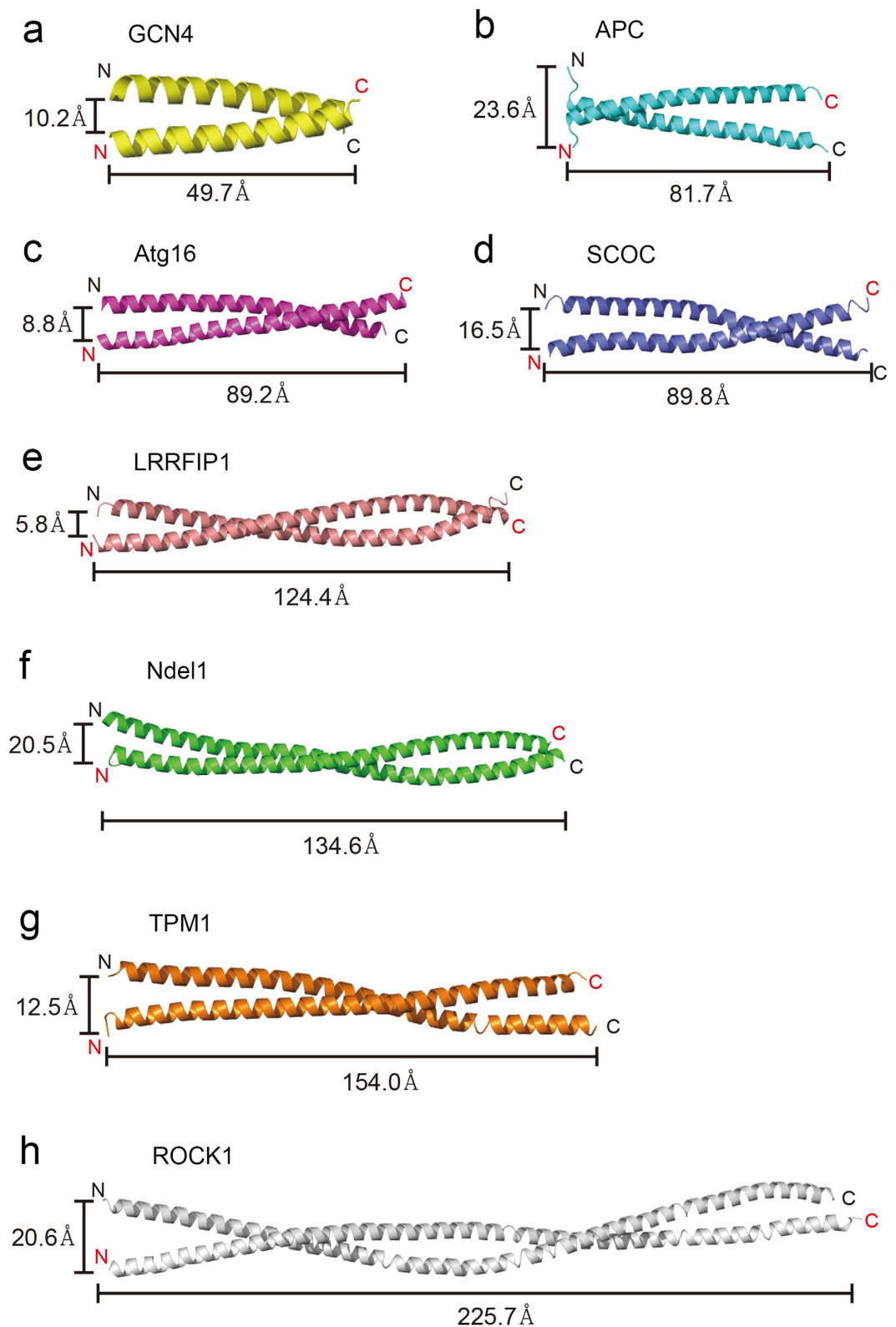
Figure 2. Structures of parallel coiled-coil domains (CCDs). (**a**) GCN4 from *Saccharomyces cerevisiae*. (**b**) APC, adenomatous polyposis coli protein from *Home sapiens*. (**c**) Atg16, autophagy protein 16 from *S. cerevisiae*. (**d**) SCOC, short coiled-coil protein from *H. sapiens*. (**e**) LRRFIP1, leucine-rich repeat flightless-interacting protein 1 from *H. sapiens*. (**f**) Ndel1, nuclear distribution protein nudE-like 1 from *Rattus norvegicus*. (**g**) TPM1, tropomyosin α-1 chain from *Gallus gallus*. (**h**) ROCK1, Rho-associated coiled-coil containing protein kinase 1 from *H. sapiens*. The length of each CCD and the Cα distance between the two N-terminal residues are provided.
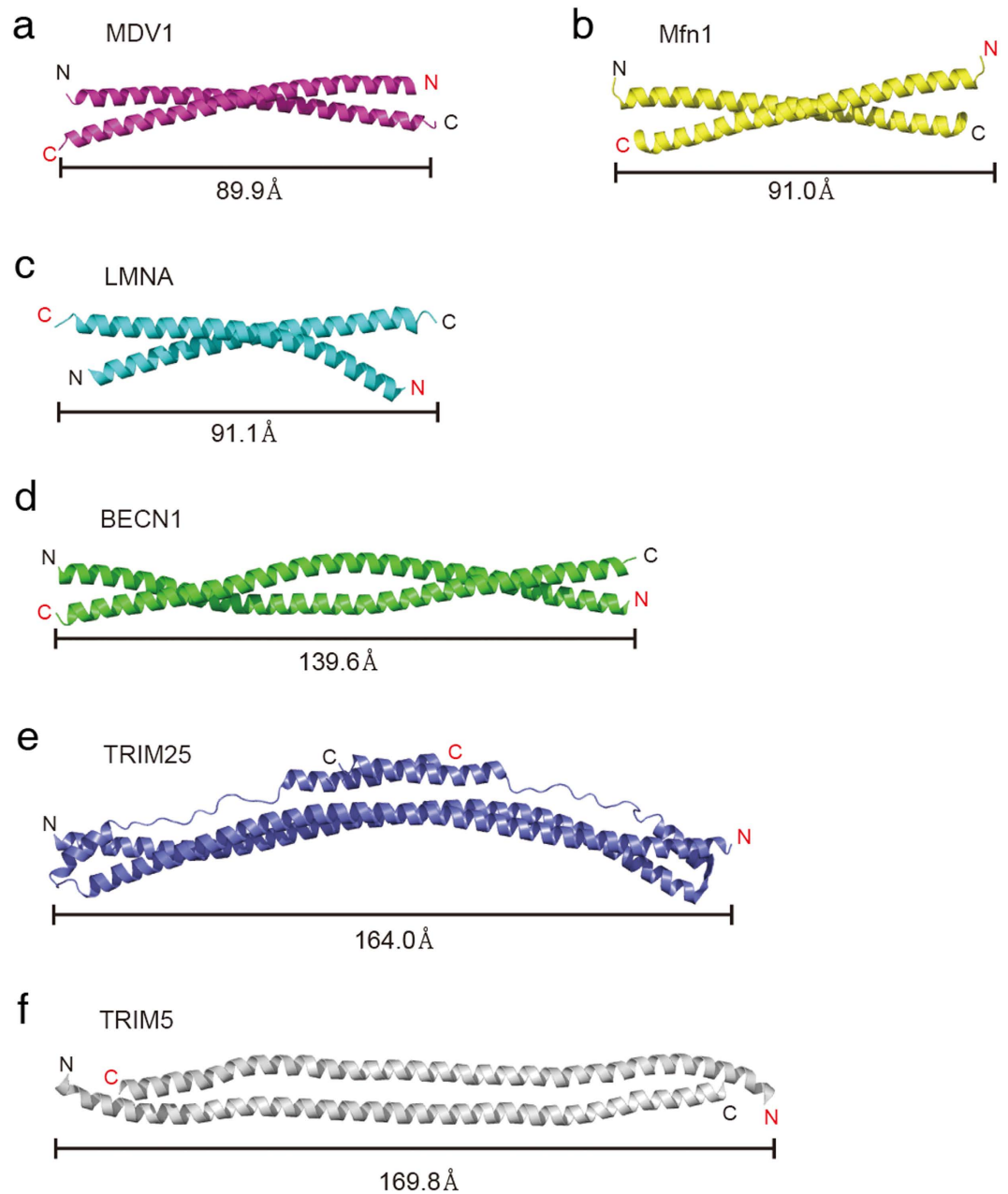
Figure 3. Structures of antiparallel coiled-coil domains (CCDs). (**a**) MDV1, mitochondrial division protein 1. (**b**) Mfn1, Mitofusin-1. (**c**) LMNA, prelamin-A/C. (**d**) BECN1, beclin-1. (**e**) TRIM25, E3 ubiquitin/ISG15 ligase tripartite motif-containing protein 25. (**f**) TRIM5, tripartite motif-containing protein 5. The length of each CCD is provided.

extremely good result compared to those of other prediction methods. LOGICOIL is very powerful for analysing the oligomeric state and orientation of CCDs[20]; however, its prediction of the orientation of dimeric CCs was not sufficiently accurate (Table 1). The first predictions were incorrect for two out of eight parallel CCDs. A bigger problem was that the prediction of antiparallel CCDs was not accurate at all. Only one out of six predictions was accurate in the case of antiparallel CCDs; the others were incorrect oligomeric structure predictions (Table 1). The antiparallel CC proteins (MDV1, Mfn1, LMNA, BECN1, TRIM25, and TRIM5; see the legend in Table 1 for the full names of these proteins) usually formed aggregates or tetramers in solution (Fig. 5a–f). The soluble tetramer is likely to be arranged as shown in Fig. 5g, but it is difficult to produce a single prediction model for higher-order oligomers (Fig. 1e). For some CCDs fused with SspB, it was difficult to obtain SEC-MALS data because the CCDs were prone to aggregation, and this phenomenon was most apparent with antiparallel CCDs. The short antiparallel CCD, MDV1, showed a significant proportion to be dimers, although the main fraction corresponded to a tetramer (Fig. 5b). Therefore, the orientation of a CCD can be easily assessed by analysing the oligomeric nature of SspB-fused CCDs. We have termed this method ACCORD.

Gene (*Origin*)	Residues	CC length (Å)	1st prediction* (score)	2nd prediction* (score)	MW of Fusion protein/CC (kDa)**	MALS (kDa)	Oligomer (ACCORD)	PDB (ref)
SspB (*E. coli*)	M1–K165	—	—	—	18.3	—	—	1OX8 (ref. 32)
Parallel (Para)								
GCN4 (*S. cerevisiae*)	R249–R281	49.7	Para dimer (1.25)	Anti dimer (0.97)	23.4/4.0	58.8	Dimer	2ZTA[7]
APC (*H. sapiens*)	A2–I55	81.7	Para dimer (1.32)	Trimer (1.19)	25.5/6.1	59.0	Dimer	1DEB[45]
Atg16 (*S. cerevisiae*)	V58–R130	89.2	Para dimer (1.37)	Anti dimer (0.89)	27.8/8.4	64.0	Dimer	3A7O[***]
SCOC (*H. sapiens*)	M78–K159	89.8	Trimer (1.11)	Anti dimer (1.03)	28.8/9.4	86.2	Dimer	4BWD[46]
LRRFIP1 (*H. sapiens*)	D162–E249	124.4	Para dimer (1.14)	Anti dimer (0.97)	30.0/10.6	71.8	Dimer	4H22 (ref. 47)
Ndel (*R. norvegicus*)	D8–D99	134.6	Para dimer (1.14)	Anti dimer (0.97)	30.4/11.0	70.0	Dimer	2V71 (ref. 48)
TPM1 (*G. gallus*)	M1–E98	154.0	Para dimer (1.05)	Tetramer (1.04)	30.6/11.2	64.6	Dimer	3U1A[49]
ROCK1 (*H. sapiens*)	L535–K709	225.7	Tetramer (1.12)	Anti dimer (1.02)	40.0/20.6	95.9	Dimer	3O0Z[34]
Antiparallel (Anti)								
MDV1 (*S. cerevisiae*)	Q231–G300	89.9	Tetramer (1.08)	Anti dimer (1.03)	27.6/8.2	65.2/107.1	Dimer/tetramer	2XU6 (ref. 43)
Mfn1 (*M. musculus*)	F660–H735	91.0	Tetramer (1.21)	Anti dimer (1.0)	28.2/8.8	90.0	Tetramer	1T3J[50]
LMNA (*H. sapiens*)	A328–S398	91.1	Tetramer (1.11)	Para dimer (1.09)	27.9/8.5	—	Aggregation	2XV5 (ref. 51)
BECN1 (*H. sapiens*)	E174–K266	139.6	Anti dimer (1.03)	Tetramer (1.02)	30.7/11.3	138.0	Tetramer	3Q8T[44]
TRIM25 (*H. sapiens*)	A194–Q356	164.0	Tetramer (1.15)	Para dimer (1.03)	38.4/18.8	119.3	Tetramer	4CFG[***]
TRIM5 (*M. mulatta*)	M133–L241	169.8	Trimer (1.16)	Tetramer (1.14)	32.6/13.2	—	Aggregation	4TN3 (ref. 52)
Prediction								
NDP52 (*H. sapiens*)	E141–E334	—	Para dimer (1.24)	Tetramer (1.06)	42.6/23.2	102.5 (2.4)	Dimer	[****]

Table 1. Summary of ACCORD results in comparison with the results of LOGICOIL prediction. Para and Anti, Parallel and Antiparallel, respectively. SspB, Stringent starvation protein B; GCN4, General control protein GCN4; APC, Adenomatous polyposis coli protein; Atg16, Autophagy protein 16; SCOC, Short coiled-coil protein; LRRFIP1, Leucine-rich repeat flightless-interacting protein 1; Ndel1, Nuclear distribution protein nudE-like 1; TPM1, Tropomyosin alpha-1 chain; ROCK1, Rho-associated coiled-coil containing protein kinase 1; MDV1, Mitochondrial division protein 1; Mfn1, Mitofusin-1; LMNA, Prelamin-A/C; BECN1, Beclin-1; TRIM25, E3 ubiquitin/ISG15 ligase tripartite motif-containing protein 25; TRIM5, Tripartite motif-containing protein 5; NDP52, Calcium-binding and coiled-coil domain-containing protein 2; *Results of coiled-coil prediction server (LOGICOIL)[20]; **SspB fusion proteins with His$_6$ tags; ***no reference available; ****no PDB ID.

Application of ACCORD to unknown CC orientations. Next, we applied the ACCORD method to CCDs with unknown orientations. NDP52, a dimeric autophagy receptor, consists of multiple domains, including a central CCD[35,36]. The CCD of NDP52 fused with SspB was subjected to SEC-MALS and was observed as a clear dimer in solution (Fig. 6a). Therefore, NDP52 was judged to be a parallel dimer based on our ACCORD method. To confirm the directionality of the CCs of NDP52, we attempted to determine the structure of the protein, but it did not crystallize successfully. Therefore, we used small angle X-ray scattering (SAXS) instead. If the fusion partner is a bigger monomeric protein, the molecular envelope is either Y-shaped (in the case of a parallel CC; Fig. 6b) or dumbbell-shaped (in the case of an antiparallel CC; Fig. 6c). The maltose binding protein (MBP) fused with the CCD of NDP52 shows a clear Y-shaped molecular envelope by SAXS (Fig. 6b), verifying that the CCD of NDP52 is parallel. In order to test an antiparallel CCD, we performed X-ray crystallography (Supplementary Fig. 2 and Supplementary Table 1) and SAXS of the MBP-MDV1 fusion protein (Supplementary Figs 3 and 4 and Supplementary Table 2). The results clearly showed that it has a dumbbell-shaped, rather than a Y-shaped, molecular envelope (Fig. 6c). These results were all consistent with those from our ACCORD method.

Discussion

Given the central role of CCs in protein structure, the ability to easily determine the orientation of the helices in a CCD is important. Undoubtedly, a method to determine the orientation of CCD helices inside a cell would be really useful. There are several approaches to determine the global topology of *E. coli* membrane proteins using fusion techniques with reporter proteins, alkaline phosphatase, and green fluorescence protein[37], and to establish the networks of basic region leucine zipper protein-protein interactions using a fluorescence resonance energy transfer-based assay[38]. These techniques are somewhat similar to ACCORD, in terms of using a fusion
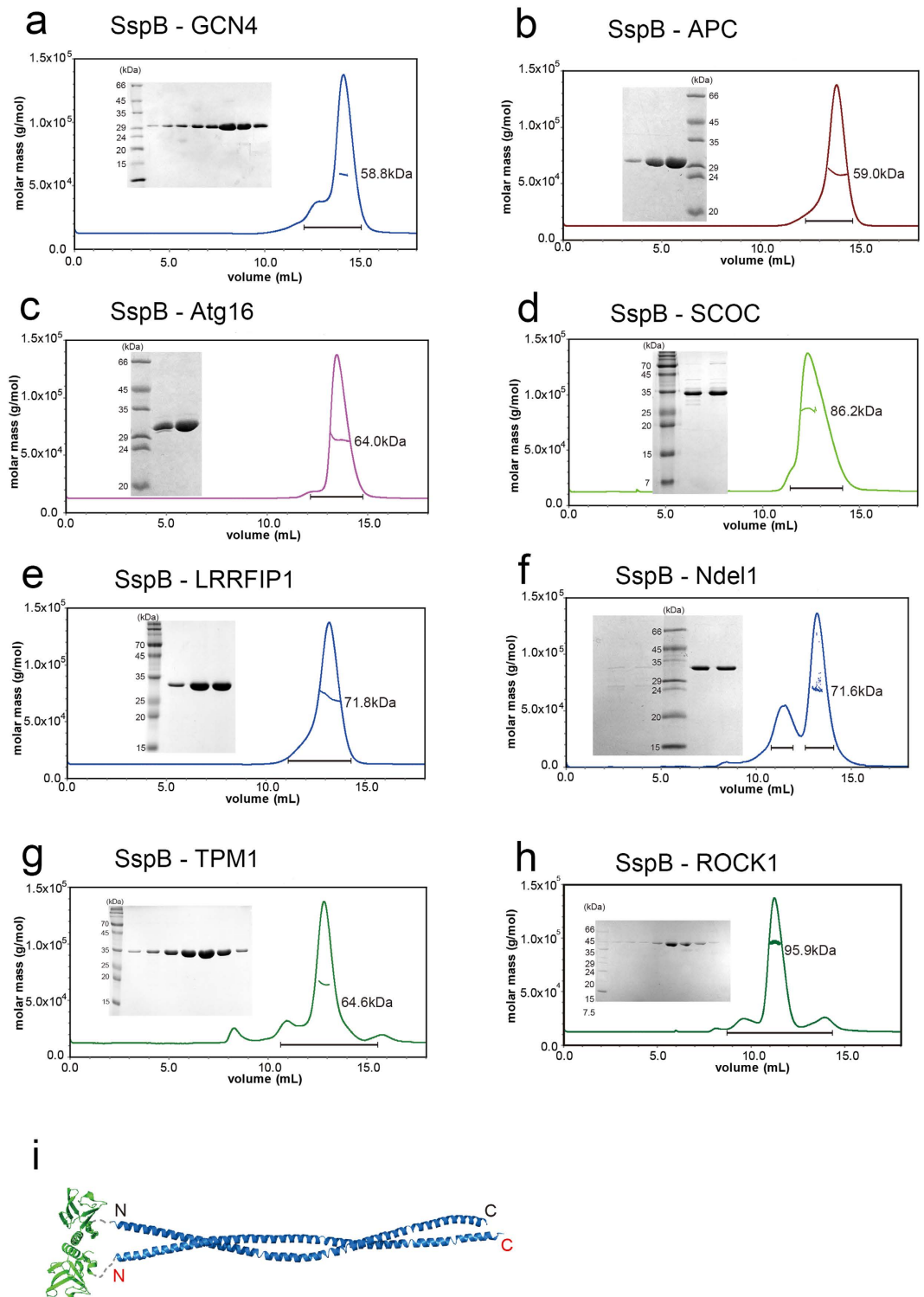
Figure 4. ACCORD results for parallel coiled-coils (CCs). Size-exclusion chromatography with multi-angle light scattering data from (**a**) SspB-GCN4, (**b**) SspB-APC, (**c**) SspB-Atg16, (**d**) SspB-SCOC, (**e**) SspB-LRRFIP1, (**f**) SspB-Ndel1, (**g**) SspB-TPM1, and (**h**) SspB-ROCK1 fusion proteins. Insets show the results of sodium dodecyl sulphate-polyacrylamide gel electrophoresis of the chromatography fractions. (**i**) Schematic model of the SspB-ROCK1 dimer. All parallel CCDs fused with the SspB protein behaved as dimers in solution.

protein; however, their purposes are completely different from that of the ACCORD technique. Furthermore, proteome-wide approaches have limitations such as false positives; thus, *in vitro* experimentation is necessary to determine the orientation of CCs in a particular protein.
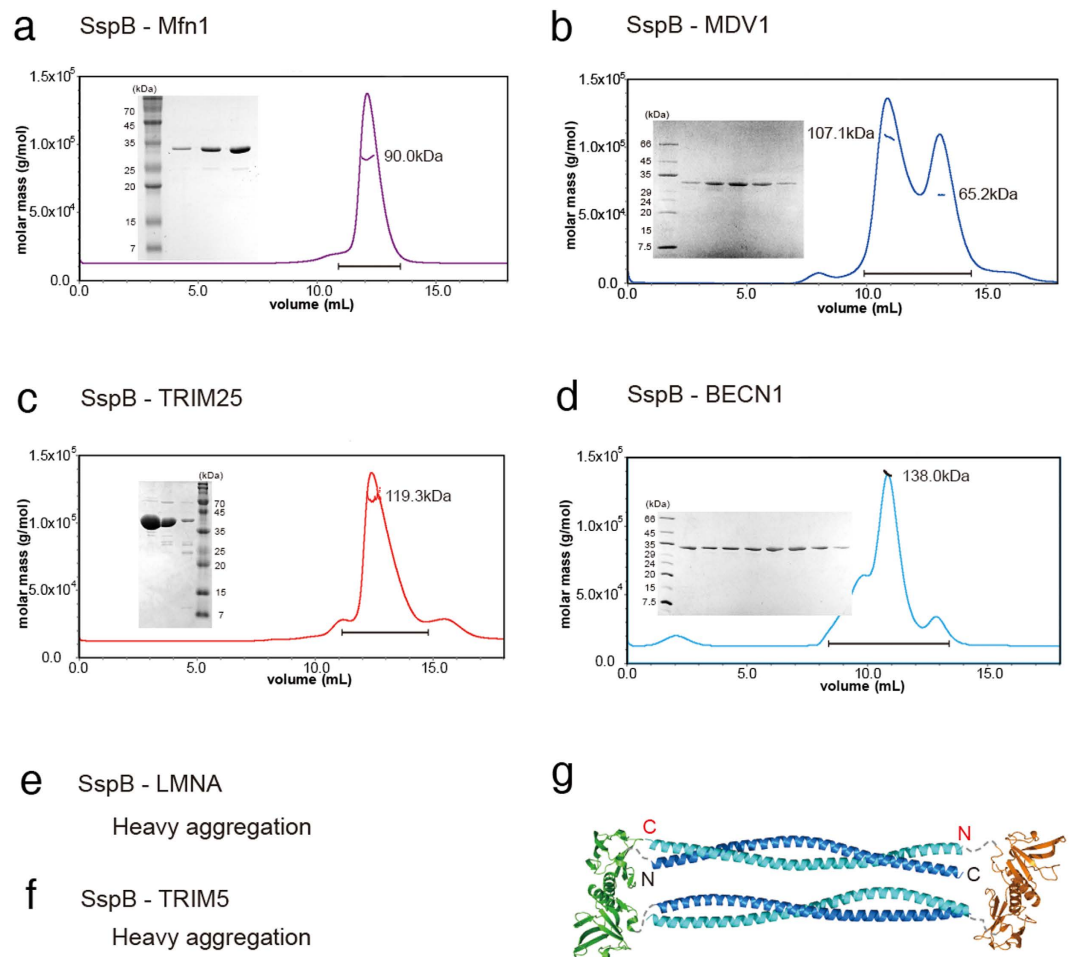
Figure 5. ACCORD results for antiparallel coiled-coils. Size-exclusion chromatography with multi-angle light scattering (SEC-MALS) data from (**a**) stringent starvation protein B (SspB)-Mfn1, (**b**) SspB-MDV1, (**c**) SspB-TRIM25, and (**d**) SspB-BECN1 fusion proteins. Insets shows the results of sodium dodecyl sulphate-polyacrylamide gel electrophoresis of the chromatography fractions. SEC-MALS analysis was not possible for (**e**) SspB-LMNA and (**f**) SspB-TRIM5 because of their significant precipitation. (**g**) Schematic model of the SspB-BECN1 tetramer. Antiparallel CCDs fused with the SspB protein behaved as tetramers or heavy aggregates in solution.

CCs are very structurally diverse and can exist in many different oligomeric states, including dimers, trimers, tetramers, and others. The majority of CCs form dimers, and the ACCORD technique was applied to homodimeric CCs because the fusion tag, SspB, is a homodimer. The use of SspB with a long C-terminal tail has pros and cons. The long and flexible tail acts as a linker that provides structural plasticity to accommodate structurally diverse CCs. However, our preliminary analysis suggests that it might be difficult to use the ACCORD technique to assess antiparallel CCs that are relatively short. For example, the ACCORD result for human MYO10 CC (51 amino acid residues and 47.1 Å; expected as a tetramer) showed that it was a dimer in solution (Supplementary Fig. 5a,b), and we speculate that it forms a dimer with the orientation shown in Supplementary Fig. 5c. Although the current version of the SspB fusion tag has some limitations, we are in the process of improving it for greater assessment accuracy.

Therefore, the ACCORD method can be used as a tool to determine homodimeric CC directionality to supplement *in silico* predictions that are inaccurate because the energy difference between the correct and incorrect directionalities of CCs is not large enough for discrimination between them. For a definitive answer, a three-dimensional structure determination by X-ray crystallography or nuclear magnetic resonance spectroscopy is best, but these methods are labour intensive and are sometimes technically infeasible because of the lack of diffracting crystals or because of the high molecular weight of a CC sample. The ACCORD approach is straightforward and is a good complement to *in silico* prediction methods to determine the CC orientation in a protein. By combining the ACCORD method and currently available prediction methods, the molecular function of target CC proteins can be interpreted with confidence.

Methods

DNA manipulation. Full-length *sspB* was amplified from *E. coli* genomic DNA by PCR and cloned into the NcoI and BamHI restriction sites of a modified pET vector (containing tobacco etch virus and thrombin cleavage
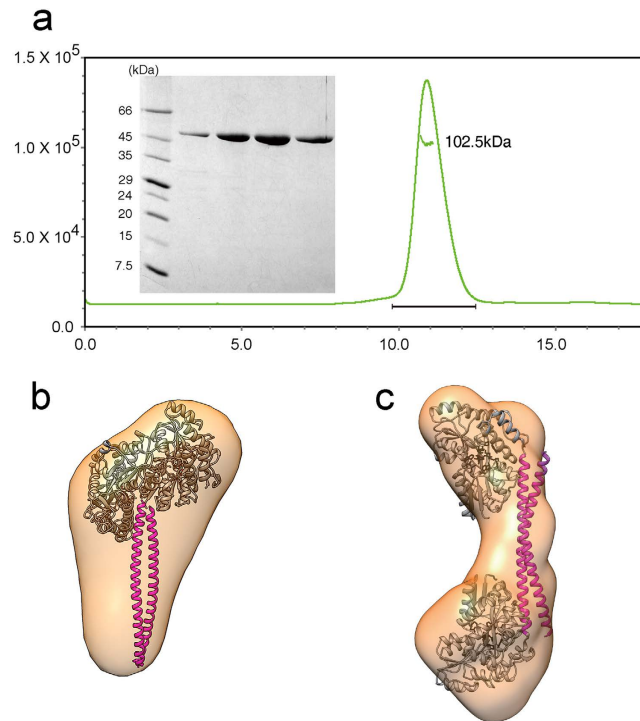
Figure 6. Validation of ACCORD results. (**a**) Size-exclusion chromatography with multi-angle light scattering (SEC-MALS) data showed that SspB-NDP52 is a dimer in solution, suggesting that the coiled-coil domain (CCD) of NDP52 is oriented in a parallel manner. Inset shows the results of SDS-PAGE of the chromatography fractions. (**b**) *Ab initio* envelope using SAXS data superimposed onto two MBP molecules (PDB ID: 3DM0) and a parallel CCD of Atg16 (PDB ID: 3A7O), which might have similar CC lengths based on the numbers of residues. (**c**) Molecular envelope of MBP-MDV1. The high-resolution crystal structure of MBP-MDV1 (Supplementary Fig. 2) was fitted into the low resolution molecular envelope generated from SAXS data.

sites) to construct an N-terminal hexa-histidine-tagged protein (hereafter, His$_6$-SspB). The gene fragments GCN4 (249–281) and MDV1 (231–300) from the genomic DNA of *Saccharomyces cerevisiae* and APC (2–55), SCOC (78–159), LRRFIP1 (162–249), Beclin-1 (174–266), ROCK1 (535–709), and NDP52 (141–334) from the cDNA of *Homo sapiens* were amplified using primers containing restriction enzyme sites (BamHI at the 5′-end and EcoRI or XhoI at the 3′-end; NEB). The gBlock gene fragments Ndel1 (8–99), MYO10 (883–933), TPM1 (1–98), Mfn1 (600–735), LMNA (328–398), TRIM25 (194–356), and TRIM5 (133–241) were synthesized using primers containing the above-mentioned restriction enzyme sites (Integrated DNA Technologies). Sequences of the CCD proteins used in this study are provided in Supplementary Table 3. Genes for the CCDs were ligated into the His$_6$-SspB vector. The gene fragments MDV1 (231–300) and NDP52 (197–270) were also ligated into a modified pMAL vector with an N-terminal hexa-histidine sequence and several point mutations to facilitate crystallization[39]. The resulting plasmids were transformed into *E. coli* BL21(DE3) cells.

Protein expression and purification. His-tagged fusion protein expression was induced by the addition of 0.5 mM isopropyl-β-D-thiogalactoside at 291 K for 20 h. Cells were harvested by centrifugation and resuspended in buffer A (50 mM Tris, pH 8.0, and 0.5 mM tris(2-carboxyethyl)phosphine) containing 300 mM NaCl. After sonication, the cell lysate was loaded onto a HisTrap column (GE Healthcare) and then eluted by gradient purification with buffer A containing 100 mM NaCl and 500 mM imidazole. Eluted proteins were concentrated by ultrafiltration (Amicon Ultra 30 K NMWL, Millipore) and loaded onto a Superdex 200 10/300 GL column (GE Healthcare) equilibrated with buffer A containing 200 mM NaCl.

Cells expressing the MBP-tagged fusion proteins were harvested by centrifugation and resuspended in buffer A containing 100 mM NaCl and 1 mM EDTA. After sonication, the cell lysate was applied to a column containing amylose resin (NEB) and collected by gravity flow. The beads were washed with 10 column volumes of buffer A, and then the protein was eluted with buffer A supplemented with 10 mM maltose. The eluted sample was further purified using a Q-FF column (GE Healthcare). Finally, the fusion proteins were loaded onto a HiLoad 16/600 Superdex 200 column (GE Healthcare) pre-equilibrated with buffer A containing 100 mM NaCl.

SEC-MALS. SEC-MALS experiments were performed using a fast protein liquid chromatography system (GE Healthcare) connected to a Wyatt MiniDAWN TREOS MALS instrument and Wyatt Optilab rEX differential refractometer. The column and buffer used were the same as those used in the final purification step. Ovalbumin was used as the isotropic scatterer for detector normalization. Light scattering from each sample (3–5 mg/ml, 0.5 ml) was measured and analysed using ASTRA V software (Wyatt).

Crystallization and structure determination. Purified MBP-MDV1 fusion protein was crystallized at 295 K using the hanging drop vapour diffusion method and mixing an equal volume of the protein and a reservoir solution containing 100 mM sodium acetate, pH 4.5, 25% (w/v) polyethylene glycol 3350, 100 mM CaCl₂, and 1.3–1.7 M sodium formate. Crystals were flash-frozen with the reservoir solution containing 20% (v/v) glycerol in a nitrogen stream at 100 K. Native MBP-MDV1 data were collected at beamline 5 C of the Pohang Accelerator Laboratory (PAL), Korea. Diffraction data were indexed, integrated, and scaled using HKL2000 software[40]. The structure was determined by molecular replacement using the MBP mutant structure as a search model[39]. Statistics for data collection and refinement are provided in Supplementary Table 1.

SAXS. Solutions of the MBP-NDP52 and MBP-MDV1 fusion proteins were prepared in buffer A containing 100 mM NaCl. The concentration of MBP-NDP52 was 2.6 mg/ml, and that of MBP-MDV1 was 9.8 mg/ml. Scattering data of MBP-NDP52 and MBP-MDV1 were collected at beamline BL-10C of the Photon Factory, Japan, and at beamline 4 C of PAL, respectively. Details of experimental parameters are shown in Supplementary Table 2. Briefly, scattering images from the proteins at various concentrations were reduced to two-dimensional (2D) data by circular integration. A preliminary analysis of these 2D data using PRIMUS (ATSAS program) provided the radius of gyration (R_g), Porod volume, and experimental molecular weight[41]. *Ab initio* modelling and averaging of these models were performed using DAMMIF and DAMAVER, respectively. Rigid body modelling of the crystallographic structure on dummy atom models was computed using the Situs program package[42].

Protein structure modeling. All protein models were generated using the crystal structures of SspB (PDB accession code: 1OX8)[32] and the CCDs of MDV1 (PDB accession code: 2XU6)[43], Beclin-1 (PDB accession code: 3Q8T)[44], Atg16 (PDB accession code: 3A7O), and ROCK1 (PDB accession code: 3O0Z)[34]. All figures for structures were generated using PyMOL (http://www.pymol.org).

References

1. Lupas, A. N. & Gruber, M. The structure of alpha-helical coiled coils. *Adv Protein Chem* **70,** 37–78 (2005).
2. Lupas, A., Van Dyke, M. & Stock, J. Predicting coiled coils from protein sequences. *Science* **252,** 1162–1164 (1991).
3. Delorenzi, M. & Speed, T. An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. *Bioinformatics* **18,** 617–625 (2002).
4. McDonnell, A. V., Jiang, T., Keating, A. E. & Berger, B. Paircoil2: improved prediction of coiled coils from sequence. *Bioinformatics* **22,** 356–358 (2006).
5. Walshaw, J. & Woolfson, D. N. Socket: a program for identifying and analysing coiled-coil motifs within protein structures. *J Mol Biol* **307,** 1427–1450 (2001).
6. Steinert, P. M. The two-chain coiled-coil molecule of native epidermal keratin intermediate filaments is a type I-type II heterodimer. *J Biol Chem* **265,** 8766–8774 (1990).
7. O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254,** 539–544 (1991).
8. Glover, J. N. & Harrison, S. C. Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature* **373,** 257–261 (1995).
9. Frank, S. *et al.* Characterization of the matrilin coiled-coil domains reveals seven novel isoforms. *J Biol Chem* **277,** 19071–19079 (2002).
10. Takeshita, K. *et al.* X-ray crystal structure of voltage-gated proton channel. *Nat Struct Mol Biol* **21,** 352–357 (2014).
11. Bochtler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartunik, H. D. & Huber, R. The structures of HsIU and the ATP-dependent protease HsIU-HsIV. *Nature* **403,** 800–805 (2000).
12. Song, H. K. *et al.* Mutational studies on HsIU and its docking mode with HsIV. *Proc Natl Acad Sci USA* **97,** 14103–14108 (2000).
13. Kim, J. H. *et al.* Insights into autophagosome maturation revealed by the structures of ATG5 with its interacting partners. *Autophagy* **11,** 75–87 (2015).
14. Freedman, S. J., Song, H. K., Xu, Y., Sun, Z. Y. & Eck, M. J. Homotetrameric structure of the SNAP-23 N-terminal coiled-coil domain. *J Biol Chem* **278,** 13462–13467 (2003).
15. Kim, J. H. & Song, H. K. Swapping of interaction partners with ATG5 for autophagosome maturation. *BMB Rep* **48,** 129–130 (2015).
16. Siegert, R., Leroux, M. R., Scheufler, C., Hartl, F. U. & Moarefi, I. Structure of the molecular chaperone prefoldin: unique interaction of multiple coiled coil tentacles with unfolded proteins. *Cell* **103,** 621–632 (2000).
17. Armstrong, C. T., Vincent, T. L., Green, P. J. & Woolfson, D. N. SCORER 2.0: an algorithm for distinguishing parallel dimeric and trimeric coiled-coil sequences. *Bioinformatics* **27,** 1908–1914 (2011).
18. Mahrenholz, C. C., Abfalter, I. G., Bodenhofer, U., Volkmer, R. & Hochreiter, S. Complex networks govern coiled-coil oligomerization–predicting and profiling by means of a machine learning approach. *Mol Cell Proteomics* **10,** M110 004994 (2011).
19. Trigg, J., Gutwin, K., Keating, A. E. & Berger, B. Multicoil2: predicting coiled coils and their oligomerization states from sequence in the twilight zone. *PLoS One* **6,** e23519 (2011).
20. Vincent, T. L., Green, P. J. & Woolfson, D. N. LOGICOIL–multi-state prediction of coiled-coil oligomeric state. *Bioinformatics* **29,** 69–76 (2013).
21. Zhou, N. E., Kay, C. M. & Hodges, R. S. Synthetic model proteins: the relative contribution of leucine residues at the nonequivalent positions of the 3-4 hydrophobic repeat to the stability of the two-stranded alpha-helical coiled-coil. *Biochemistry* **31,** 5739–5746 (1992).
22. Steinmetz, M. O. *et al.* A distinct 14 residue site triggers coiled-coil formation in cortexillin I. *EMBO J* **17,** 1883–1891 (1998).
23. Grigoryan, G., Reinke, A. W. & Keating, A. E. Design of protein-interaction specificity gives selective bZIP-binding peptides. *Nature* **458,** 859–864 (2009).
24. Steinkruger, J. D., Woolfson, D. N. & Gellman, S. H. Side-chain pairing preferences in the parallel coiled-coil dimer motif: insight on ion pairing between core and flanking sites. *J Am Chem Soc* **132,** 7586–7588 (2010).
25. Gayen, S. & Gruber, G. Disulfide linkage in the coiled-coil domain of subunit H of A1AO ATP synthase from Methanocaldococcus jannaschii and the NMR structure of the C-terminal segment H(85-104). *FEBS Lett* **584,** 713–718 (2010).
26. Pandya, M. J., Cerasoli, E., Joseph, A., Stoneman, R. G., Waite, E. & Woolfson, D. N. Sequence and structural duality: designing peptides to adopt two stable conformations. *J Am Chem Soc* **126,** 17016–17024 (2004).
27. Hillar, A., Tripet, B., Zoetewey, D., Wood, J. M., Hodges, R. S. & Boggs, J. M. Detection of alpha-helical coiled-coil dimer formation by spin-labeled synthetic peptides: a model parallel coiled-coil peptide and the antiparallel coiled coil formed by a replica of the ProP C-terminus. *Biochemistry* **42,** 15170–15178 (2003).

28. Wuo, M. G., Mahon, A. B. & Arora, P. S. An Effective Strategy for Stabilizing Minimal Coiled Coil Mimetics. *J Am Chem Soc* **137,** 11618–11621 (2015).
29. Levchenko, I., Seidel, M., Sauer, R. T. & Baker, T. A. A specificity-enhancing factor for the ClpXP degradation machine. *Science* **289,** 2354–2356 (2000).
30. Park, E. Y., Lee, B. G., Hong, S. B., Kim, H. W., Jeon, H. & Song, H. K. Structural basis of SspB-tail recognition by the zinc binding domain of ClpX. *J Mol Biol* **367,** 514–526 (2007).
31. Dougan, D. A., Weber-Ban, E. & Bukau, B. Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA+ protein ClpX. *Mol Cell* **12,** 373–380 (2003).
32. Song, H. K. & Eck, M. J. Structural basis of degradation signal recognition by SspB, a specificity-enhancing factor for the ClpXP proteolytic machine. *Mol Cell* **12,** 75–86 (2003).
33. Wah, D. A., Levchenko, I., Rieckhof, G. E., Bolon, D. N., Baker, T. A. & Sauer, R. T. Flexible linkers leash the substrate binding domain of SspB to a peptide module that stabilizes delivery complexes with the AAA+ ClpXP protease. *Mol Cell* **12,** 355–363 (2003).
34. Tu, D. *et al.* Crystal structure of a coiled-coil domain from human ROCK I. *PLoS One* **6,** e18080 (2011).
35. Kim, B. W., Hong, S. B., Kim, J. H., Kwon do, H. & Song, H. K. Structural basis for recognition of autophagic receptor NDP52 by the sugar receptor galectin-8. *Nat Commun* **4,** 1613 (2013).
36. Kim, B. W., Kwon do, H. & Song, H. K. Structure biology of selective autophagy receptors. *BMB Rep* **49,** 73–80 (2016).
37. Daley, D. O., Rapp, M., Granseth, E., Melen, K., Drew, D. & von Heijne, G. Global topology analysis of the Escherichia coli inner membrane proteome. *Science* **308,** 1321–1323 (2005).
38. Reinke, A. W., Baek, J., Ashenberg, O. & Keating, A. E. Networks of bZIP protein-protein interactions diversified over a billion years of evolution. *Science* **340,** 730–734 (2013).
39. Moon, A. F., Mueller, G. A., Zhong, X. & Pedersen, L. C. A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction. *Protein Sci* **19,** 901–913 (2010).
40. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. In: *Methods Enzymol.* (eds Carter J, C. W., Sweet, R. M.). Academic Press (1997).
41. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *Journal of Applied Crystallography* **45,** 342–350 (2012).
42. Wriggers, W. & Birmanns, S. Using Situs for flexible and rigid-body fitting of multiresolution single-molecule data. *Journal of Structural Biology* **133,** 193–202 (2001).
43. Koirala, S. *et al.* Molecular architecture of a dynamin adaptor: implications for assembly of mitochondrial fission complexes. *J Cell Biol* **191,** 1127–1139 (2010).
44. Li, X. *et al.* Imperfect interface of Beclin1 coiled-coil domain regulates homodimer and heterodimer formation with Atg14L and UVRAG. *Nat Commun* **3,** 662 (2012).
45. Day, C. L. & Alber, T. Crystal structure of the amino-terminal coiled-coil domain of the APC tumor suppressor. *J Mol Biol* **301,** 147–156 (2000).
46. Behrens, C., Binotti, B., Schmidt, C., Robinson, C. V., Chua, J. J. & Kuhnel, K. Crystal structure of the human short coiled coil protein and insights into SCOC-FEZ1 complex formation. *PLoS One* **8,** e76355 (2013).
47. Nguyen, J. B. & Modis, Y. Crystal structure of the dimeric coiled-coil domain of the cytosolic nucleic acid sensor LRRFIP1. *J Struct Biol* **181,** 82–88 (2013).
48. Derewenda, U. *et al.* The structure of the coiled-coil domain of Ndel1 and the basis of its interaction with Lis1, the causal protein of Miller-Dieker lissencephaly. *Structure* **15,** 1467–1481 (2007).
49. Rao, J. N., Rivera-Santiago, R., Li, X. E., Lehman, W. & Dominguez, R. Structural analysis of smooth muscle tropomyosin alpha and beta isoforms. *J Biol Chem* **287,** 3165–3174 (2012).
50. Koshiba, T., Detmer, S. A., Kaiser, J. T., Chen, H., McCaffery, J. M. & Chan, D. C. Structural basis of mitochondrial tethering by mitofusin complexes. *Science* **305,** 858–862 (2004).
51. Kapinos, L. E., Burkhard, P., Herrmann, H., Aebi, U. & Strelkov, S. V. Simultaneous formation of right- and left-handed anti-parallel coiled-coil interfaces by a coil2 fragment of human lamin A. *J Mol Biol* **408,** 135–146 (2011).
52. Goldstone, D. C. *et al.* Structural studies of postentry restriction factors reveal antiparallel dimers that enable avid binding to the HIV-1 capsid lattice. *Proc Natl Acad Sci USA* **111,** 9609–9614 (2014).

Acknowledgements

Author Contributions

B.W.K., M.K.K., D.H.K., S.H.P., J.H.K., Y.B.K., S.J.O., L.K., B.H.K., and W.S.Y. prepared the samples and performed the biochemical experiments; B.W.K. and H.K.S. designed the experiments and analysed the data; B.W.K. determined the X-ray structure of MBP-MDV1; Y.O.J. determined the SAXS solution structure. B.W.K. and H.K.S wrote the manuscript.

Additional Information

Accession code: Atomic coordinates and structure factor files for MBP-MDV1 have been deposited in the Protein Data Bank with the accession code 5JST.

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Kim, B.-W. *et al.* ACCORD: an assessment tool to determine the orientation of homodimeric coiled-coils. *Sci. Rep.* **7,** 43318; doi: 10.1038/srep43318 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.