Crystal Structure of the Cytoskeleton-associated Protein Glycine-rich (CAP-Gly) Domain

Received for publication, August 20, 2002
Published, JBC Papers in Press, September 7, 2002, DOI 10.1074/jbc.M208512200

Songlin Li‡, Jim Finley‡, Zhi-Jie Liu‡, Shi-Hong Qiu‡, Hongli Chen‡, Chi-Hao Luan‡, Mike Carson‡, Jun Tsao‡, David Johnson‡, Guangda Lin‡, Jun Zhao‡, Willie Thomas‡, Lisa A. Nagy‡, Bingdong Sha‡, Lawrence J. DeLucas‡, Bi-Cheng Wang‡, and Ming Luo‡‡**

From the Southeast Collaboratory for Structural Genomics, University of Georgia, Athens, Georgia 30602, the ‡Center for Biophysical Sciences and Engineering, ‡Department of Microbiology, and ‡Department of Cell Biology, Southeast Collaboratory for Structural Genomics, University of Alabama at Birmingham, Birmingham, Alabama 35294 and the §Department of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia 30602

Cytoskeleton-associated proteins (CAPs) are involved in the organization of microtubules and transportation of vesicles and organelles along the cytoskeletal network. A conserved motif, CAP-Gly, has been identified in a number of CAPs, including CLIP-170 and dynactins. The crystal structure of the CAP-Gly domain of Caenorhabditis elegans F53F4.3 protein, solved by single wavelength sulfur-anomalous phasing, revealed a novel protein fold containing three β-sheets. The most conserved sequence, GKNGD, is located in two consecutive sharp turns on the surface, forming the entrance to a groove. Residues in the groove are highly conserved as measured from the information content of the aligned sequences. The C-terminal tail of another molecule in the crystal is bound in this groove.

The cytoskeleton maintains the hierarchy of a eukaryotic cell to control the spatial distribution and the movement of vesicles, organelles, and large protein complexes. It mainly involves three types of protein filaments: microtubules, intermediate filaments, and actin filaments. There are extensive interactions between the filaments mediated by a large number of cytoskeleton-associated proteins (1, 2). For instance, microtubules and actin filaments are cross-linked by microtubule-associated proteins such as myosin-kinesin complexes, myosin-CLIP-170 complexes, and dynein-dynactin complexes (1, 3). These protein molecules or molecular assemblies have one end that binds to the microtubule and the other end that binds to different vesicles or organelles. Movements related to nucleus positioning or mitotic spindle orientation are orchestrated between microtubules and actin filaments through dynamic interactions of different cytoskeleton-associated proteins.

The cytoskeleton is a dynamic network that disassembles and reassembles constantly according to functional requirements. Regulation of the dynamic assembly of the cytoskeleton and motor proteins involves other non-motor accessory proteins. The term "tIPS" has recently been proposed to describe microtubule plus-end-tracking proteins (4). This class of proteins tends to localize to the plus end of microtubules, and these proteins are the most conserved microtubule-associated proteins. One common motif, named glycine-rich cytoskeleton-associated protein (CAP-Gly) domain, was first discovered in restin (or CLIP-170), the prototype +TIP, and three other proteins (5). This domain is present in a P fam family of 69 proteins (6). The amino acid sequence of the CAP-Gly domain is conserved from the simple single cell eukaryote Saccharomyces cerevisiae to the human being. The functions of this family of proteins are mostly hypothetical. They are grouped together because of the existence of this CAP-Gly domain. Therefore, the CAP-Gly domain must have a signature role in this protein family. The CAP-Gly domain can appear once or multiple times in one protein, either in the N terminus or the C terminus. It can also be found in some motor-associated proteins, such as dyactin chain 1 (7) and Drosophila kinesin-73 (8). The two CAP-Gly domains in CLIP-170 located at the N terminus were shown to mediate the binding of the CLIP-170 protein to the microtubules (9). The C-terminal region of CLIP-170 carries the domains for organelle binding, suggesting that CLIP-170 is a microtubule-organelle linker protein (10). The CAP-Gly domains of CLIP-170 bind specifically to the growing (plus) end of the microtubule (11); whereas the opposite end of CLIP-170 mediates interactions with dynactin (12) or myosin (13). This characteristic feature might allow CLIP proteins to participate in the dynamic control of cargo transport along microtubules or between microtubules and actin filaments.

The genome of Caenorhabditis elegans has been completely sequenced, and several proteins were predicted to have a CAP-Gly domain based on homology. The genes from C. elegans have been selected as targets by the Southeast Collaboratory for Structural Genomics (14). Here we report the structure solution of the CAP-Gly domain of the C. elegans F53F4.3 protein with a novel fold, which provided an example for how high throughput x-ray crystallography might impact on our knowledge of new genes identified from genome sequencing.

EXPERIMENTAL PROCEDURES

Protein Preparation and Crystallization—Based on the annotation of the C. elegans genome performed by the C. elegans sequencing consortium (15), primers were designed for cloning all predicted cDNAs into an ENTRY vector in the GATEWAY conveying system (Invitrogen) (16). Our group then took these ENTRY vectors and systematically recombinated them with an expression vector, pDEST 17.1, in a 96-well format.

1 The abbreviations used are: CAP, cytoskeleton-associated protein; CAP-Gly, glycine-rich CAP; MES, 4-morpholineethanesulfonic acid; ISAS, iterative single wavelength anomalous signal.
Recombinant protein expression of each vector was screened, and vectors that expressed soluble proteins in *Escherichia coli* were selected. F53F4.3 was one of the selected vectors that produced a high amount of soluble protein when protein expression in *E. coli* was induced at 18°C. This expression vector produced a protein that has a hexahistidine tag and an 8-amino acid peptide in front of the N terminus and an 8-amino acid peptide following the C terminus of the gene. The His tag was included for purification by a nickel affinity column, and both 8-amino acid peptides at each terminus were the result of the recombination reaction during cloning. After nickel affinity, fast protein liquid chromatography gel filtration (S-75, Amersham Biosciences), and ion exchange (Resource Q, Amersham Biosciences) chromatography purification, the purified protein was concentrated to about 10 mg/ml for crystal screening in a robotic system. Before any crystals were produced, it was noticed that the purified F53F4.3 protein had been cleaved into two halves in the solution stored at 4°C. The fragments were isolated from the SDS-PAGE gel and subjected to mass spectrometry analysis for their molecular weights and automated N-terminal amino acid sequencing. The C-terminal fragment corresponds to residues 121–229 and was subsequently cloned into a pET 28b vector (Novagen) for protein expression. The resulting protein with a hexahistidine tag, a thrombin cleavage site in front of the N terminus of the gene. The His tag by thrombin cleavage did not appear to affect crystallization. Diffraction data were collected with crystals prepared using proteins without the His tag.

**X-ray Diffraction Data**—Data for phasing were collected to 2.5 Å resolution on a single crystal and flash-cooled to 100 K at beamline ID-17 (IMCA-CAT), Advanced Photon Source (APS), Argonne National Laboratory using a MarResearch 165-mm CCD detector and 1.74 Å X-rays following the protocol described in Liu et al. (17). HKL2000 (18) was used to determine a data collection strategy (99% completion), and data collection was divided into four passes. The first and third passes cover one anomalous diffraction wedge (86–146°), whereas the second and fourth passes cover the other anomalous diffraction wedge (286–326°). The oscillation angle was 1° per frame for all the passes. Data processing was carried out using HKL2000 (18). The processed data from different passes were then merged. Data collection and processing statistics are given in Table I.

**Phase Determination**—Low resolution data (3.5 Å) were used to locate the sulfur atom sites. A total of 170 Bijvoet reflection pairs greater than 2σ, between 20.0 and 3.5-Å resolution (Bijvoet difference R-factor = 1.5%), were used to determine three sulfur atom positions (see Fig. 3A) with SOLVE V1.16 (19). These three sulfur sites could also be easily determined by the program SHELX (20). PHASES (21) was then used to refine the sulfur positions and to locate one additional sulfur site by Bijvoet difference Fourier analysis. These heavy atom sites correspond to sulfur atoms of two cysteines, a sulfur atom of one methionine, and a possible Cl– ion in the solvent region based on the final refined protein structure. All heavy atom sites were treated as sulfur for the phase calculations.

The four sulfur sites were used to resolve the enantiomorphous space groups P6122 and P6522. The handedness test feature in ISAS2001 (22) was used at 20.0–3.5-Å resolution to identify P6,22 as the correct space group because of the high figure-of-merit and low map-inversion R-value. The same four sulfur sites were used to estimate the protein phases at 3.0-Å resolution using iterative single wavelength anomalous signal (ISAS) (22). After three filters and two cycles of phase extension, the overall average figure-of-merit and map-inversion R-values were 0.73 and 0.281, respectively. The original electron density map is shown in Fig. 3B, in which the polypeptide chain could be traced for about 65 residues. After a poly-Ala model was fitted into the density, 36 residues and the calculated phases from the poly-Ala model were provided to wARP program (23) with diffraction data between 50–1.8 Å collected on a Raxis IV detector mounted on a Rigaku RU300 generator. A new poly-Ala model with 95 residues was generated. Sequence alignment and refinement were carried out with the programs O (24) and XPLOR3.5 (25). The final model of 109 residues was refined against a data set to 1.77 Å resolution collected at APS (see Table I).

**RESULTS**

**Structure Determination**—Crystals of the wild type recombinant F53F4.3 protein (residues 121–292) were grown in conditions found in a robotic screen using the 96-well sitting drop plate plates. The optimal reservoir solution consisted of 1.8 M ammonium sulfate and 0.1 M MES buffer at pH 6.5. The protein drop contained 5 μl of the reservoir solution and 5 μl of the protein solution at 10 mg/ml concentration in 20 mM HEPES buffer, pH 7.4. Crystals of 0.4 × 0.4 × 0.3 mm in size appeared in the drops in 4 days at room temperature. The space group was determined to be P6,22 with a = b = 64.16 Å, c = 101.95 Å by x-ray diffraction data collected on a Raxis-IV image plate detector (MSC). To obtain phases for solving the crystal structure, the ISAS method was employed (22). It has been argued that the anomalous signal from the sulfur atoms present in the wild type proteins would be sufficient for phase determination using only single wavelength x-ray data (17). To collect x-ray diffraction data with high accuracy required for the ISAS method, crystals of F53F4.3 were cryo-frozen at 100 °C, and x-ray data were collected at the wavelength of 1.74 Å to increase the anomalous signal from sulfur atoms while minimizing crystal damage. After screening a number of crystals, a data set to 2.5-Å resolution was completed from the best crystal. This data set to the moderate resolution limit was collected for phasing only. Three sulfur atom positions could be located by the SOLVE program using SAS data. The Patterson peaks corresponding to the sulfur sites were clearly visible (Fig. 1A). A fourth site was found in further refinement. Phases derived...
Refinement statistics could be defined as three antiparallel sheets. The C-terminal 6 residues (Leu-224 to Ile-229) protrude the two-stranded sheet is on top of the second three-stranded extended continuously to the first strand of the second sheet. L-shaped to each other with the last strand of the first sheet ordered part of the N terminus has a 9-residue H9251/sequence conservation (Fig. 3A). This rigorous statistical procedure compares each position in the given alignment with the observed distribution of each amino acid in the C. elegans genome. A variety of structural alphabets may be employed in this analysis, for example, grouping the amino acids by charge or chemical class. The data presented here treat each of the 20 amino acids independently. A portion of the alignment and its annotation are shown in Fig. 3B.

The CAP-Gly domain appears to be a rigidly packed motif in the middle of two extended polypeptides. The structure of this domain is likely to represent a new protein fold. A search for three-dimensional homology (27) found no similar structures from the non-redundant subset of the Protein Data Bank.

Conserved Features—The sequence of the F53F4.3 domain was provided to the EBI server (www.ebi.ac.uk/fasta33) to search the Swiss Protein Database using the FASTA alignment method (28). The 58 aligned sequences, ranging from 55 to 22% identity, corresponded to the CAP-Gly motif (5, 29). The Shannon information content (30) is used as the measure of sequence conservation (Fig. 3A). This rigorous statistical procedure compares each position in the given alignment with the probability of randomly distributed amino acids weighted by the observed distribution of each amino acid in the C. elegans genome. A variety of structural alphabets may be employed in this analysis, for example, grouping the amino acids by charge or chemical class. The data presented here treat each of the 20 amino acids independently. A portion of the alignment and its annotation are shown in Fig. 3B.

The longest stretch of highly conserved sequence corresponds to Gly-189 to Gly-193, which is located at the surface of the domain between β3 and β4. This 5-residue GKNKD segment is conserved in most homologous sequences except in a few sequences where the asparagine is substituted by a histidine or the aspartic acid is substituted by a serine. A loop structure opposite to this conservative stretch, Val-156 to Met-160, has higher average temperature factors, indicating that this part of the structure may be flexible. This is a highly variable region among the homologous sequences, which could present a characteristic recognition site for other protein interactions if the conserved stretch is to associate with the cytoskeleton. There is a conserved aromatic cluster including Tyr-168, Phe-174, Trp-179, Tyr-200, Phe-201, and Phe-210. The hydrophobic cluster may represent a stabilizing core structure. There is also a significant groove adjacent to the conserved 189–193 loop. The groove is lined by 12 residues: Val-156, Gln-159, Arg-162, Phe-174, Trp-179, Tyr-184, Lys-190, Asn-191, Val-195, Gly-209, Phe-210, and Val-211.

The ordered part of the N terminus has a 9-residue α-helix (Asp-135–Lys-144) with relatively higher average temperature factors. This α-helix is preceded by 17 disordered residues that do not make any stable contacts with the rest of this domain and could be available for intermolecular interactions. The three-layer β/β structure contains seven antiparallel β-strands. The first layer consists of β1, β2a, and β7. The second layer starts with the extension of β-strand 2 (β2b) and continues with β3 and β6. The last layer contains β4 and β5. The topology of the CAP-Gly domain could be defined as three antiparallel β-sheets of 3-3-2 strands. The two three-stranded sheets are L-shaped to each other with the last strand of the first sheet extended continuously to the first strand of the second sheet. The two-stranded sheet is on top of the second three-stranded sheet. The C-terminal 6 residues (Leu-224 to Ile-229) protrude out of the globular domain. Considering the disordered residues at the N terminus prior to the α-helix, the CAP-Gly domain appears to be a rigidly packed motif in the middle of two extended polypeptides. The structure of this domain is likely to represent a new protein fold. A search for three-dimensional homology (27) found no similar structures from the non-redundant subset of the Protein Data Bank.

Overall Structure—The C-terminal domain of F53F4.3 is roughly spherical with a three-layer β/β structure (Fig. 2). The ordered part of the N terminus has a 9-residue α-helix (Asp-135–Lys-144) with relatively higher average temperature factors. This α-helix is preceded by 17 disordered residues that do not make any stable contacts with the rest of this domain and could be available for intermolecular interactions. The three-layer β/β structure contains seven antiparallel β-strands. The first layer consists of β1, β2a, and β7. The second layer starts with the extension of β-strand 2 (β2b) and continues with β3 and β6. The last layer contains β4 and β5. The topology of the CAP-Gly domain could be defined as three antiparallel β-sheets of 3-3-2 strands. The two three-stranded sheets are L-shaped to each other with the last strand of the first sheet extended continuously to the first strand of the second sheet. The two-stranded sheet is on top of the second three-stranded sheet. The C-terminal 6 residues (Leu-224 to Ile-229) protrude out of the globular domain. Considering the disordered residues at the N terminus prior to the α-helix, the CAP-Gly domain appears to be a rigidly packed motif in the middle of two extended polypeptides. The structure of this domain is likely to represent a new protein fold. A search for three-dimensional homology (27) found no similar structures from the non-redundant subset of the Protein Data Bank.

The CAP-Gly domain appears to be a rigidly packed motif in the middle of two extended polypeptides. The structure of this domain is likely to represent a new protein fold. A search for three-dimensional homology (27) found no similar structures from the non-redundant subset of the Protein Data Bank.

Conserved Features—The sequence of the F53F4.3 domain was provided to the EBI server (www.ebi.ac.uk/fasta33) to search the Swiss Protein Database using the FASTA alignment method (28). The 58 aligned sequences, ranging from 55 to 22% identity, corresponded to the CAP-Gly motif (5, 29). The Shannon information content (30) is used as the measure of sequence conservation (Fig. 3A). This rigorous statistical procedure compares each position in the given alignment with the probability of randomly distributed amino acids weighted by the observed distribution of each amino acid in the C. elegans genome. A variety of structural alphabets may be employed in this analysis, for example, grouping the amino acids by charge or chemical class. The data presented here treat each of the 20 amino acids independently. A portion of the alignment and its annotation are shown in Fig. 3B.

The longest stretch of highly conserved sequence corresponds to Gly-189 to Gly-193, which is located at the surface of the domain between β3 and β4. This 5-residue GKNKD segment is conserved in most homologous sequences except in a few sequences where the asparagine is substituted by a histidine or the aspartic acid is substituted by a serine. A loop structure opposite to this conservative stretch, Val-156 to Met-160, has higher average temperature factors, indicating that this part of the structure may be flexible. This is a highly variable region among the homologous sequences, which could present a characteristic recognition site for other protein interactions if the conserved stretch is to associate with the cytoskeleton. There is a conserved aromatic cluster including Tyr-168, Phe-174, Trp-179, Tyr-200, Phe-201, and Phe-210. The hydrophobic cluster may represent a stabilizing core structure. There is also a significant groove adjacent to the conserved 189–193 loop. The groove is lined by 12 residues: Val-156, Gln-159, Arg-162, Phe-174, Trp-179, Tyr-184, Lys-190, Asn-191, Val-195, Gly-209, Phe-210, and Val-211. Some of the hydrophobic residues are part of the conserved hydrophobic cluster, such as Phe-174, Trp-179, and Phe-210. A number of their main chain and side chain atoms are solvent-accessible. In the crystal lattice, the extended C terminus of one molecule

Table I X-ray crystallography data statistics

| Crystal | Crystallization         | 1.8 M (NH₄)₂SO₄, 0.1 M Mes, pH 6.5 |
|         | Space group            | P6₁,2₂ |
|         | Cell parameters (Å)    | a = 64.156, b = 64.156, c = 101.946 |

Data collection and processing statistics

| Data set used for phasing | Wavelength (Å) | 1.74 |
|                          | Resolution (Å) | 1.77 |
|                          | Reflections    | 12,651 |
|                          | Completeness (%) | 99.8 |
|                          | Bijvoet Redundancy | 16.9 |
|                          | Rmerge (%)     | 3.2 (6.3 in the last shell) |

Data set used for refinement

| Reflections used in refinement | Wavelength (Å) | 1.54 |
|                               | Resolution (Å) | 1.77 |
|                               | Reflections    | 12,651 |
|                               | Completeness (%) | 99.2 (95.0 in the last shell) |
|                               | Rmerge (%)     | 9.8 |
|                               | Lσ (I)         | 19.0 (2.9 in the last shell) |

Refinement statistics

| Reflections used in refinement | Water molecules | 88 |
|                               | Residues modeled | 135–229 |
|                               | Ramachandran plot | 84.4% core region |
|                               | Average B-factors for protein | 28.816 |
|                               | Errors (not including glycine residues) | 38.593 |
|                               | R.m.s.d. of bond length (Å) | 0.021 |
|                               | R.m.s.d. of bond angles (°) | 1.70 |

* R = Σ|Fobs| - |Fcal|/Σ|Fobs|.

** Highest resolution of data set followed by resolution range in highest resolution shell.

| Crystallization | 1.8 M (NH₄)₂SO₄, 0.1 M Mes, pH 6.5 |
|-----------------|----------------------------------|
| P6₁,2₂          |                                  |
| a = 64.156      | b = 64.156                       |
| c = 101.946     |                                  |

Fig. 2. The CAP-Gly structure. Left, a schematic topology drawing of the CAP-Gly domain of C. elegans F53F4.3 protein. The β-strands are represented by arrows. There are three β-sheets formed by three, three, and two strands, respectively. The strand of β2a and β2b is continuous. Right, a ribbon drawing of the three-dimensional structure of the CAP-Gly domain. The color coding is the same as the topology drawing.
is packed into the groove of a symmetry-related molecule. The side chain of residue Glu-228 forms a buried salt bridge with Arg-162 and a hydrogen bond with residue Tyr-184. Residue 162 or its neighbor is usually basic according to the sequence alignment. Hydrogen bonds are also formed between the main chain atoms of residue Ile-229 and the symmetry-related molecule. Lys-190 is also close to the acidic C terminus.

The CAP-Gly domains contain a large number of conserved glycine residues. Two of them, Gly-193 and Gly-197, are involved in the formation of two consecutive sharp turns with the second one as a type II $\beta$-turn. This unique double-turn motif exposes the most conserved region (GKNDG) in the CAP-Gly domain. The residue after the GKNDG sequence is often a serine/threonine, which forms a hydrogen bond with the side chain of residue Asp-192. Other conserved glycine residues include: Gly-170 and Gly-177, which are involved in bending $\beta 2 b$ and $\beta 3$, resulting in the loop connecting $\beta 2 b$ and $\beta 3$ to be perpendicular; Gly-189, which is in the middle of an extended loop containing the most conserved stretch; Gly-181, which is packed closely to the conserved Phe-210 (3.5 Å); and Gly-208 to Phe-201 (4 Å). These conserved glycine residues may be required for maintaining the folding of the three-layer sheet.
DISCUSSION

The crystal structure of a widely distributed CAP-Gly domain is reported to high resolution. This domain was isolated from the C. elegans F53F4.3 protein when expressed in E. coli. Based on previous sequence alignment, the CAP-Gly domain was proposed to contain 52 amino acids (Pam: PFOI302) (5). These 52 amino acids constitute the conserved core region of this domain. However, the crystal structure revealed that this domain consists of 84 amino acids in this protein. One α-helix and three β-sheets form a novel protein fold not observed previously. Two conserved regions are identified on the surface based on information content analysis of the amino acid sequences from this protein family. A surface loop containing residues 193–197 (GRKNO) could present itself for interaction with the microtubule. The second feature is a groove that was occupied in our crystal structure by the C terminus of a neighboring molecule. A helix at the N terminus is loosely formed in the multiplex protein structure. The C terminus is usually located in the middle of a long polypeptide chain. The C terminus could represent hypothetically the binding peptide residues in this region could readily be available to fit into a disordered region. It is possible that a helix would only be present in high B-factor in this structure and is preceded by a long disordered region. The specificity for the binding location could be provided by sequence variations of the nonconserved regions, or interacting with other regions. The CAP-Gly domain has also been found in dynemin/dynactin (P150Glued) complexes. The conventional view is that the motor proteins could get on and off the microtubule by alternated motor domain binding to the filament. The function of the CAP-Gly domain in this mobile complex may be to act as a tether to “glue” the complex on the microtubule. This will keep the traveling complex on the right track without totally falling off the microtubule.

We demonstrated here that protein crystal structures could be determined in an HTP manner using the naturally present sulfur sites as anomalous scatterers provided that high quality crystals could be grown. The prerequisite for growing high quality protein crystals is the production of highly purified, rigid, globular protein molecules. It seems inevitable that subcloning of digested fragments from soluble protein preparations is frequently employed to identify the portion of a gene product that will ultimately be crystallized. Sequence analysis of protein homology may not predict the correct structural domain that forms crystals. In this case, the predicted domain was much smaller than the one that was suitable for crystallization. When such a strategy is implemented, novel protein structure folds that are correlated to significant biological functions might be solved without prior sequence-based target selection. Since each protein family classified by sequence homology contains many members, it is difficult to predict which creative member could yield new crystals for structure determination. Shotgun screening in an HTP manner is an alternative approach to complement other target selection strategies.

Acknowledgments—We thank the Mass Spectrometry and Protein Sequencing facilities at the University of Alabama at Birmingham (UB) Comprehensive Cancer Center for their services. We thank John Rose, Albert Fu, Jeff Habel, Florian Schubert, and Wolfgang Tempel for assistance in analysis of anomalous diffraction data.

REFERENCES

1. Goede, B. L., Drubin, D. G., and Barnes, G. (2000) Curr. Opin. Cell Biol. 12, 68–74.
2. Houseweart, M. K., and Cleveland, D. W. (1999) Curr. Biol. 9, R864–R866.
3. Fuchs, E., and Yang, Y. (1999) Cell 98, 547–550.
4. Schuyler, S. C., and Pellman, D. (2001) Cell 105, 421–424.
5. Riehemann, K., and Sorg, C. (1993) Trends Biochem. Sci. 18, 82–83.
6. Bateman, A., Birney, E., Cerruti, L., Durbin, R., Ettlinger, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M., and Sonnhammer, E. L. (2002) Nucleic Acids Res. 30, 376–380.
7. Collin, G. B., Nishina, P. M., Marshall, J. D., and Naggett, J. K. (1998) Genomics 53, 359–364.
8. Liu, H., Liu, Z. M., and Nirenberg, M. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 1086–1091.
9. Pierre, F., Schoel, J., Rickard, J. E., and Kreis, T. E. (1999) Cell 99, 547–550.
10. Schoel, J., Pieper, P., Rickard, J. E., Diamantopoulos, G. S., Valetti, C., van der Goot, F. G., Haner, M., Aebi, U., and Kreis, T. E. (1999) J. Biol. Chem. 274, 25883–25891.
11. Peres, F., Diamantopoulos, G. S., Stalder, R., and Kreis, T. E. (1999) Cell 96, 517–527.
12. Vaughan, K. T., Tynan, S. H., Faulkner, N. E., Echeverri, C. J., and Vallele, R. B. (1999) J. Cell Sci. 112, 1437–1447.
The CAP-Gly Structure

48601

13. Lantz, V. A., and Miller, K. G. (1998) J. Cell Biol. 140, 897–910
14. Norvell, J. C., and Machalek, A. Z. (2000) Nat. Struct. Biol. 7, (suppl.), 931–931
15. The C. E. S. Consortium (1998) Science 282, 2012–2018
16. Wallout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., and Vidal, M. (2000) Methods Enzymol. 328, 575–592
17. Liu, Z. J., Vysotski, E. S., Chen, C. J., Rose, J. P., Lee, J., and Wang, B. C. (2000) Protein Sci. 9, 2088–2093
18. Owinski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
19. Terwilliger, T. C. (1997) Methods Enzymol. 276, 530–537
20. Sheldrick, G. M. (1990) Acta Crystallogr. Sect. A 46, 463–466
21. Furey, W., and Swaminathan, S. (1997) Methods Enzymol. 277, 590–620
22. Wang, B. C. (1985) Methods Enzymol. 115, 90–112
23. Lamzin, V. S., and Perrakis, A. (2000) Nat. Struct. Biol. 7, (suppl.), 978–81
24. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–9
25. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–21
26. Brünger, A. T. (1992) X-FLOR Version 3.1: A System for X-ray Crystallography and NMR, Yale University Press, New Haven and London
27. Gibart, J.-F., Madej, T., and Bryant, S. H. (1996) Curr. Opin. Struct. Biol. 6, 377–385
28. Pearson, W. R. (1990) Methods Enzymol. 183, 63–98
29. Watanabe, T. K., Shimizu, F., Nogata, M., Kawai, A., Fujiiwara, T., Nakamura, Y., Takahashi, E., and Hirai, Y. (1996) Cytogenet. Cell Genet. 72, 208–211
30. Karlin, S., and Altschul, S. F. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 2264–2268
31. Bilbe, G., Delahie, J., Bruggen, J., Richener, H., Asselbergs, F. A., Cerletti, N., Serg, C., Odink, K., Tarsay, L., Wiesendanger, W., et al. (1992) EMBO J. 11, 2103–2113
32. Biggins, G. R., Warren, W., Seal, S., Takahashi, M., Rapley, E., Barfoot, R. H., G., Brown, C., Biggs, P. J., Lakhani, S. R., Jones, C., Hansen, J., Blair, E., Hofmann, B., Siebert, R., Turner, G., Evans, D. G., Schrander-Stumpel, C., Beemer, F. A., van den Ouweland, A., Hildey, D., Dolpe, B., Cleveland, M. G., Leigh, I., Leisti, J., Rasmussen, S., Wallace, M. R., Fenlake, C., Banerjee, P., Oiso, N., Chaggar, R., Merrett, S., Leonard, N., Huber, M., Hohl, D., Chapman, P., Burn, J., Swift, S., Smith, A., Ashworth, A., and Stratton, M. R. (2000) Nat. Genet. 25, 160–165
33. Laiti, J. H., Lee, B. M., and Wright, P. E. (2001) Curr. Opin. Struct. Biol. 11, 39–46
34. Carson, M. (1997) Methods Enzymol. 277, 493–505