

A Spring-Loaded State of NusG in Its Functional Cycle Is Suggested by X-ray Crystallography and Supported by Site-Directed Mutants[‡]

J. Randy Knowlton,[§] Mikhail Bubunenkov,^{||} Michelle Andrykovitch,[§] Wei Guo,^{§,⊥} Karen M. Routzahn,[§] David S. Waugh,[§] Donald L. Court,^{||} and Xinhua Ji^{*,§}

Macromolecular Crystallography Laboratory and Gene Regulation and Chromosome Biology Laboratory,
National Cancer Institute, P.O. Box B, Frederick, Maryland 21702

Received November 27, 2002; Revised Manuscript Received January 18, 2003

ABSTRACT: Transcription factor NusG is present in all prokaryotes, and orthologous proteins have also been identified in yeast and humans. NusG contains a 27-residue KOW motif, found in ribosomal protein L24 where it interacts with rRNA. NusG in *Escherichia coli* (EcNusG) is an essential protein and functions as a regulator of Rho-dependent transcription termination, phage λ N and rRNA transcription antitermination, and phage HK022 Nun termination. Relative to EcNusG, *Aquifex aeolicus* NusG (AaNusG) and several other bacterial NusG proteins contain a variable insertion sequence of ~70 residues in the central region of the molecule. Recently, crystal structures of AaNusG in space groups $P2_1$ and $I222$ have been reported; the authors conclude that there are no conserved dimers among the contacting molecules in the crystals [Steiner, T., Kaiser, J. T., Marinkovic, S., Huber, R., and Wahl, M. C. (2002) *EMBO J.* 21, 4641–4653]. We have independently determined the structures of AaNusG also in two crystal forms, $P2_1$ and $C222_1$, and surprisingly found that AaNusG molecules form domain-swapped dimers in both crystals. Additionally, polymerization is also observed in the $P2_1$ crystal. A unique “ball-and-socket” junction dominates the intermolecular interactions within both oligomers. We believe that this interaction is a clue to the function of the molecule and propose a spring-loaded state in the functional cycle of NusG. The importance of the ball-and-socket junction for the function of NusG is supported by the functional analysis of site-directed mutants.

The elongation activity of RNA polymerase (RNAP)¹ is controlled in part by the action of competing regulatory complexes. For instance, transcription termination activity in *Escherichia coli* competes with the antitermination activity of a multiprotein complex comprised of the bacteriophage λ N protein and N-utilization substances NusA, NusB, NusE, and NusG (1–3). NusG facilitates the termination activity of Rho protein both in vivo and in vitro (4, 5). NusG binds Rho (5) and is the bridge between Rho and RNAP (6). In fact, NusG only stably binds the *E. coli* RNAP elongation complex in the presence of Rho (6). Rho binds the nascent RNA during elongation, using the RNA to reach polymerase and cause termination at specific sites. NusG modulates this

termination activity through RNAP. In this reaction, no specific or direct binding of NusG to the RNA has been demonstrated. To date, there is no evidence that any bacterial NusG forms oligomers in solution.

Although it was initially identified as a component of the λ N antitermination complex, *E. coli* NusG (EcNusG) was subsequently shown to be a regulator of Rho-dependent transcription termination (7–12), phage HK022 Nun termination (4, 13), and rRNA transcription antitermination (3–5, 14–18). NusG contains a KOW motif, a 27-residue sequence (Figure 1) that is also found in ribosomal protein L24 (16), where it is involved in the binding of L24 to 23S rRNA.

NusG is present in all prokaryotes, and orthologous proteins have been identified in yeast (19) and humans (20). *Aquifex aeolicus* NusG (AaNusG) and several other bacterial NusG proteins contain an insertion sequence of ~70 residues that is not present in EcNusG (Figure 1). Recently, Steiner et al. (21) reported the crystal structures of AaNusG in two space groups, $P2_1$ and $I222$. Four crystallographically independent molecules (Mol A–D) exist in the $P2_1$ structure, and one molecule (Mol E) is in $I222$. Therefore, the structures of five independent molecules were determined, revealing that AaNusG is comprised of three discrete domains (D1–D3) with an overall appearance of loosely tethered disks lying in a plane. D1 and D3 are derived from conserved sequences

[‡] Atomic coordinates and structure factors have been deposited with the Protein Data Bank as entries 1NPP and 1NPR for the $P2_1$ and $C222_1$ crystals, respectively.

* To whom correspondence should be addressed: NCI-Frederick, 1050 Boyles St., Bldg. 539, Frederick, MD 21702. Phone: (301) 846-5035. Fax: (301) 846-6073. E-mail: jix@ncifcrf.gov.

[§] Macromolecular Crystallography Laboratory.

^{||} Gene Regulation and Chromosome Biology Laboratory.

[⊥] Present address: Center for Veterinary Medicine, Food and Drug Administration, Rockville, MD 20855.

¹ Abbreviations: AaNusG, *A. aeolicus* NusG; Mol A–E, molecules A–E, respectively; D1–D3, domains 1–3, respectively; EcNusG, *E. coli* NusG; MAD, multiwavelength anomalous diffraction; RNAP, RNA polymerase; SeMet, selenomethionyl.

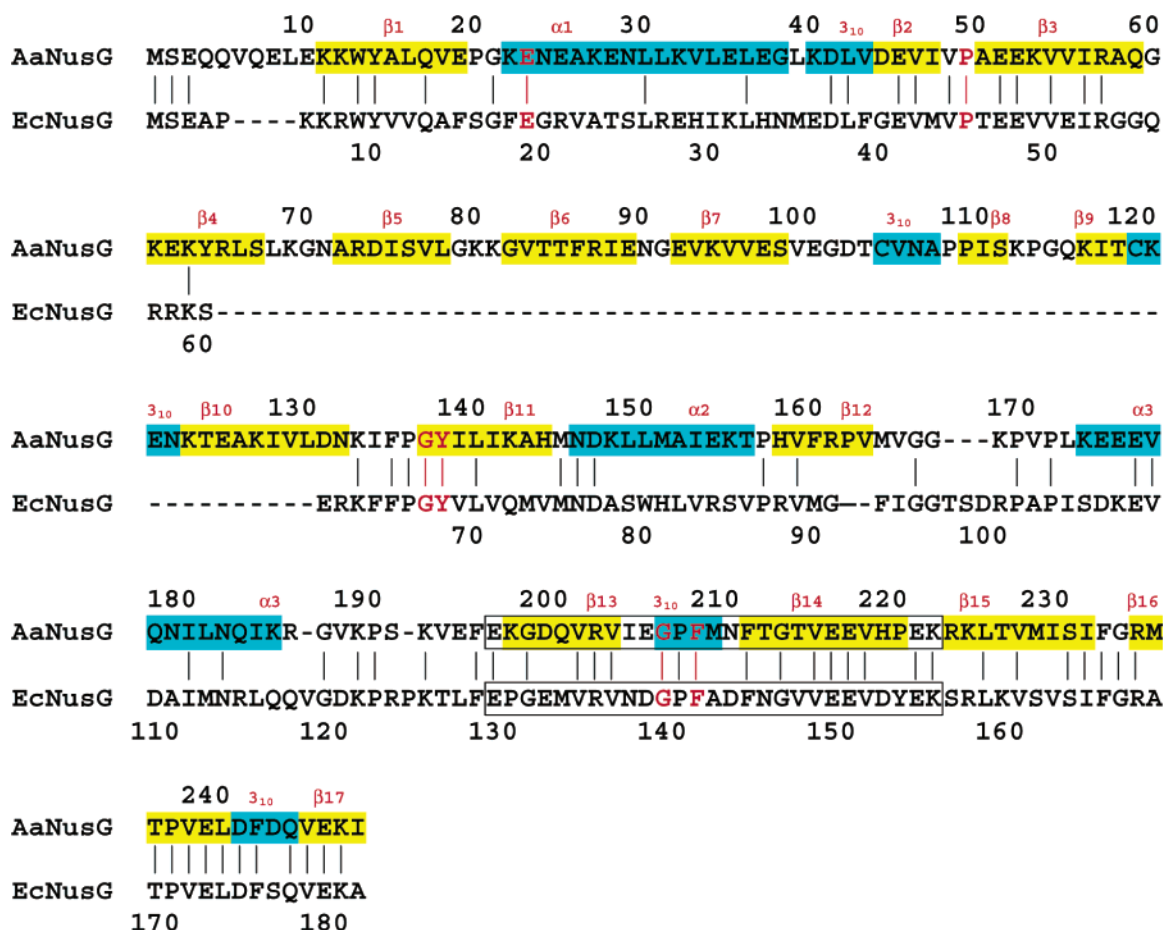


FIGURE 1: Sequence alignment of AaNusG and EcNusG. Residue numbers are shown above for AaNusG and below for EcNusG. Strictly conserved residues are indicated by vertical bars; residues strictly conserved in 30 sequences are in red. A blue background indicates helices and yellow β -strands. The boxed sequence is the KOW motif (16).

with counterparts in all NusG proteins. D2, on the other hand, is composed of residues corresponding to the variable insertion sequence (Figure 1). Mol A–D contain all three domains, whereas Mol E does not contain D3, although a gel analysis of dissolved crystals showed the full-length molecule. A common feature of the five molecules is the flexible N-terminus, resulting in unobserved stretches of four to nine residues. Another highly flexible region is a six-residue linker between D1 and D3. Among the five molecules, this flexible linker is observed in Mol A only. With regard to the intermolecular interactions and crystal packing of AaNusG, Steiner and co-workers conclude that there are no conserved dimers among the contacting molecules in the crystals (21). Here, we report the crystal structures of AaNusG also in two space groups ($P2_1$ and $C222_1$) with five independent molecules (Mol A–D in the $P2_1$ structure and Mol E in the $C222_1$ crystal). Our $P2_1$ crystal may be the same as the $P2_1$ structure previously reported, based on their unit cell dimensions (2.8% difference in b and $<0.5\%$ in a , c , and β); our $C222_1$ crystal is a new form. Moreover, each of our five molecules contains all three domains as well as the six-residue flexible linker between D1 and D3. Surprisingly, we have observed that AaNusG form domain-swapped dimers in both $P2_1$ and $C222_1$ crystals. Additionally, polymerization is also observed in the $P2_1$ crystal. Strikingly, a unique “ball-and-socket” junction dominates the intermolecular interaction in both oligomeric states. We believe that this interaction is a clue about the function of the molecule

and propose a spring-loaded state in the functional cycle of NusG. The importance of the ball-and-socket junction for the function of NusG is supported by the functional analysis of site-directed mutants.

EXPERIMENTAL PROCEDURES

Expression, Purification, and Crystallization of AaNusG. Cloning, expression, and purification of the native protein and crystallization of $C222_1$ crystals have been reported previously (22). The selenomethionyl (SeMet) protein was prepared by the methionine pathway inhibition method (23). A single colony of *E. coli* BL21(DE3) containing a maltose-binding protein–AaNusG fusion plasmid was used to inoculate 100 mL of Luria broth (LB). Approximately 15 mL of a culture grown overnight was used to inoculate 1000 mL of the same medium. The culture was grown at 37 °C until it reached an OD_{600} of 0.5–1.0, at which point the cells were pelleted by centrifugation. The cell pellet was washed with SeMet medium and centrifuged again. The cells were resuspended in 1000 mL of SeMet medium, allowed to grow for 10–15 min at 37 °C, and then induced with 1 mM isopropyl β -D-thiogalactopyranoside (IPTG) and 100 ng/mL anhydrotetracycline hydrochloride. At the time of induction, the temperature was reduced to 30 °C and the cells were grown overnight, harvested by centrifugation, and stored at -80 °C. The purification procedure for SeMet AaNusG was the same as that for the native protein (22). Purified AaNusG was concentrated with an Amicon stirred cell unit. The final

Table 1: X-ray Diffraction Data Statistics for AaNusG Crystals

| | MAD ($P2_1$) | | | native | |
|--|-------------------------|-------------------------|---------------------------|---------|----------|
| | λ_{edge} | λ_{peak} | λ_{remote} | $P2_1$ | $C222_1$ |
| wavelength (Å) | 0.9792 | 0.9788 | 0.9636 | 1.0019 | 1.5418 |
| unit cell | | | | | |
| a (Å) | 87.19 | 87.15 | 87.22 | 87.06 | 65.95 |
| b (Å) | 55.39 | 55.34 | 55.40 | 55.86 | 124.58 |
| c (Å) | 112.64 | 112.51 | 112.69 | 112.61 | 83.60 |
| α and γ (deg) | 90 | 90 | 90 | 90 | 90 |
| β (deg) | 90.55 | 90.57 | 90.55 | 90.39 | 90 |
| resolution range (Å) | 20–2.20 | 20–2.20 | 20–2.20 | 20–2.00 | 35–2.22 |
| redundancy | 3.11 | 4.13 | 3.20 | 3.11 | 4.11 |
| overall | 96.2 | 92.9 | 95.6 | 98.8 | 96.3 |
| completeness (%) | | | | | |
| last shell | 53.2 | 40.5 | 82.0 | 92.2 | 85.2 |
| completeness (%) ^a | | | | | |
| overall $I/\sigma(I)$ | 14.83 | 10.49 | 9.05 | 16.83 | 29.66 |
| last shell $I/\sigma(I)$ | 2.06 | 1.74 | 1.44 | 1.88 | 1.87 |
| overall R_{scaling} ^b | 0.092 | 0.080 | 0.079 | 0.067 | 0.058 |
| last shell R_{scaling} ^b | 0.419 | 0.386 | 0.403 | 0.538 | 0.619 |

^a Ranges of 2.24–2.20, 2.07–2.00, and 2.26–2.22 Å for MAD, native $P2_1$, and $C222_1$ crystals, respectively. ^b $R_{\text{scaling}} = \sum |I - \langle I \rangle| / \sum I$.

concentration of the protein was 32 mg/mL in 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 2 mM EDTA, and 10 mM DTT.

Crystals of SeMet AaNusG were grown by the hanging drop vapor diffusion method. The drops, containing 2.5 μ L of protein solution and 2.5 μ L of reservoir solution, were microseeded with native AaNusG crystals and equilibrated with 500 μ L of reservoir solution. The reservoir solution was derived from the Hampton Cryo Screen [0.085 M NaHepes (pH 7.5), 15% glycerol, 8.5% 2-propanol, and 17% PEG 4000], with the addition of 1% dioxane. The crystals grew at 18 ± 1 °C and reached dimensions of 0.2 mm \times 0.1 mm \times 0.05 mm in 5–7 days. The native crystals in space group $P2_1$ were grown under the same conditions, but in the absence of dioxane, and took 10–14 days to reach a similar size.

Size Exclusion Chromatography. To determine the oligomeric state of AaNusG in solution, size exclusion chromatography was performed at 4 and 20 °C using Superdex 75 resin packed in an XK 26/10 column (Pharmacia). All experiments were performed in phosphate-buffered saline (PBS) (pH 7.2) at a protein concentration of 1.8 mg/mL (64.3 μ M). The molecular mass standards RNase A (13.7 kDa), ovalbumin (43.0 kDa), and albumin (67.0 kDa) were from the Amersham Pharmacia LMW Gel Filtration Calibration Kit (lot 189982). The average retention volume for each protein was measured at both 4 and 20 °C. The final results (not shown) were the average of three to five repetitive measurements.

Data Collection, Phasing, and Refinement. We determined the $P2_1$ structure by multiwavelength anomalous diffraction (MAD). Single-wavelength data for a native crystal and MAD data for a SeMet variant were collected at beamline X9B of the National Synchrotron Light Source, Brookhaven National Laboratory (Upton, NY), with an ADSC Quantum 4 CCD detector. The crystals were flash-frozen in liquid nitrogen and maintained at -173 °C with an Oxford Cryosystem. On the basis of the fluorescence scan, three energies near the Se K-absorption edge corresponding to the edge, peak, and remote wavelengths were selected (Table 1). X-ray data were processed using HKL2000 (24). There

Table 2: Refinement Statistics for AaNusG Native Structures

| | $P2_1$ | $C222_1$ |
|---|---------|----------|
| resolution range (Å) | 20–2.00 | 35–2.22 |
| data used for refinement | 72021 | 15493 |
| data used for R_{free} calculations | 5485 | 749 |
| no. of residues | 961 | 242 |
| no. of water oxygen atoms | 858 | 146 |
| no. of 2-propanol molecules | 2 | 0 |
| final R -factor ^a | 22.0 | 23.6 |
| final R -free | 27.8 | 27.4 |
| rms ^b deviations from ideal geometry | | |
| bond distances (Å) | 0.012 | 0.008 |
| bond angles (deg) | 1.7 | 1.4 |
| Ramachandran plot | | |
| most favored ϕ and ψ angles (%) | 83.7 | 83.3 |
| disallowed ϕ and ψ angles (%) | 0.0 | 0.0 |

^a Crystallographic R -factor = $\sum_{hkl} ||F_o| - |F_c|| / \sum_{hkl} |F_o|$. ^b Root-mean-square.

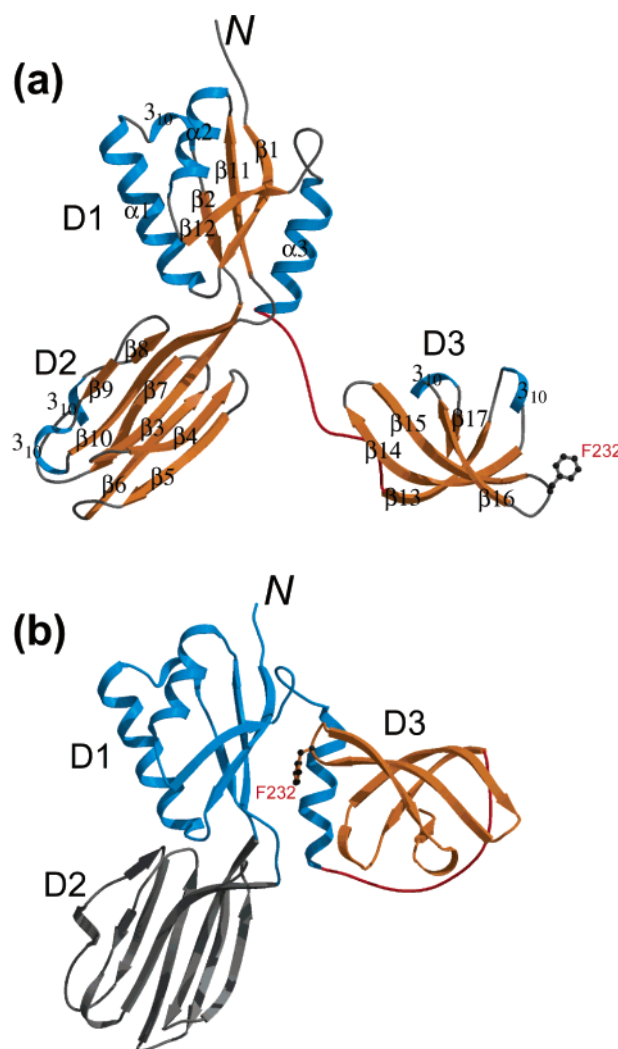


FIGURE 2: Overall structure of AaNusG. (a) A schematic representation of Mol A with three domains. Helices are shown as blue spirals, β -strands as orange arrows, and loops as gray tubes with the flexible linker in red. (b) A proposed molecular model of the spring-loaded state. D1 is blue, D2 gray, D3 orange, and the flexible linker red. F232 is highlighted. This figure was prepared with Molscript (39) and Raster3D (40).

are four AaNusG molecules in the asymmetric unit, corresponding to a solvent content of $\sim 45\%$. A total of 24 Se sites, including one with two alternate positions, were located with SOLVE (25). The solution had a figure of merit of 0.46.

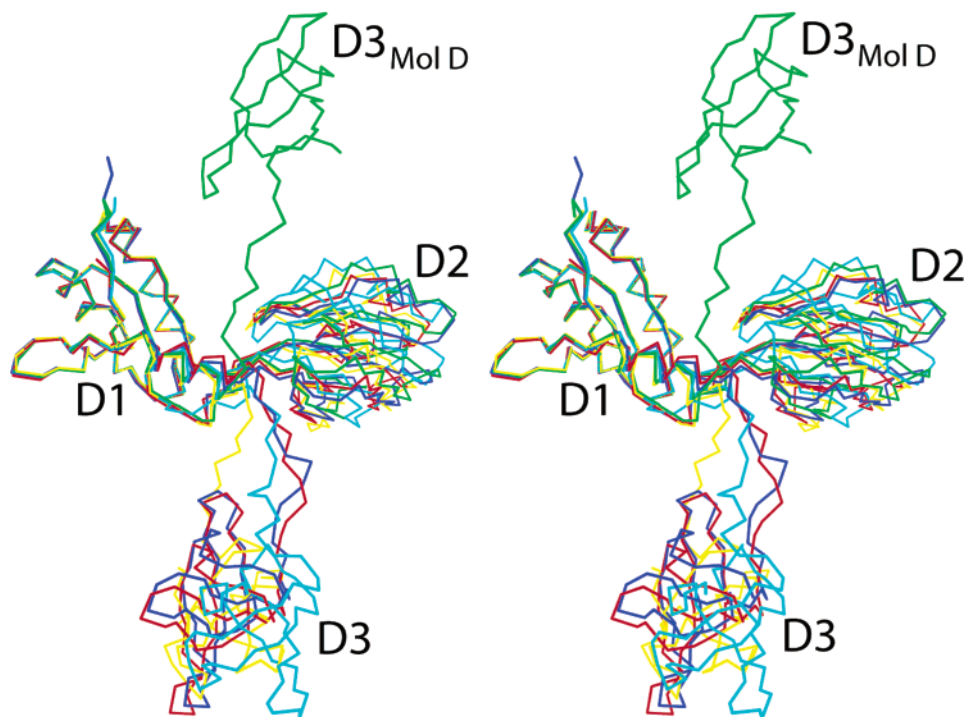


FIGURE 3: Conformational flexibility of AaNusG in stereo. Mol A (blue), B (yellow), C (red), D (green), and E (cyan) are aligned on the basis of the optimized superposition of C α atoms in D1. This figure was prepared with Molscript (39).

The Se positions were refined with SHARP (26), and the phases were improved with DM (27) and Solomon (28). ARP/WARP (29) was used to refine and extend the phases to 2.14 Å and to build the initial model containing 560 of 996 residues. On the basis of the initial model and the improved phases from ARP/WARP (29), a total of 848 residues were built, including a continuous chain of 234 residues within one monomer. Multiple rounds of refinement were performed using CNS (30) with 7.6% of the data reserved for cross validation. As the refinement progressed, the main chain atoms of flexible loops were traced into electron density maps calculated with phases from Solomon (28) phase recombination. The refinement was coupled with bulk solvent correction. Model building was carried out with O (31). The quality of the structure was assessed using PROCHECK (32). The refinement statistics can be found in Table 2.

The X-ray diffraction data of the C222₁ crystal were collected and processed as previously described (22), but with improved resolution (Table 1). The C222₁ structure was determined by molecular replacement using AmoRe (33) in combination with difference Fourier synthesis. Mol A of the P2₁ structure was used as the starting model. Molecular replacement trials starting with the whole molecule and with the combination of D1 and D2 did not result in any meaningful solution. Using D1 only, however, resulted in an outstanding solution. Difference Fourier synthesis based on D1 located D2; D3 was then positioned by the difference Fourier synthesis with a model containing both D1 and D2. The structure was refined using the same procedure as for the P2₁ crystal, and the statistics can be found in Table 2.

Cloning of *E. coli* nusG by Gap Repair Recombination. The *E. coli* nusG gene was cloned by retrieval to a plasmid from the chromosome using the gap repair (34, 35) version of recombineering (36). The pCR-Blunt plasmid (Invitrogen) was amplified by PCR with primers bearing the 40-nucleotide

extensions homologous to the very 5'- and 3'-coding regions of nusG (forward primer, CCGGTAGAGCTGGACTTCAGC-CAGGTTGAAAAAGCCTAACCTGAATTCTGCAGATAT-CCATCAC; and reverse primer, CAGGCCGCGCCCGCT-TCCCTTGAGCTTCGGTATTTCGACTCATAGCTGTTT-CCTGTGTGAAATTG). This way, the cloned nusG open reading frame should start directly from the AUG initiation codon of the lacZ gene presented on the pCR plasmid, thereby acquiring all regulatory elements of the lacZ gene. Recombineering was carried out in *E. coli* strain DY330 according to a standard procedure (37). The nusG-containing pCR plasmids were identified by restriction analysis of the insert. One plasmid clone, pAB90, was chosen, and the integrity of its nusG gene was verified by DNA sequencing. The NusG protein was expressed from this vector in *E. coli* XL1 cells.

Site-Directed Mutagenesis of *E. coli* nusG and Functional Analysis of the Mutants. The *E. coli* nusG gene in pAB90 was mutagenized using the QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA) according to the manufacturer's protocol. The ball element F165 was replaced with either A, D, T, V, or Y. Residue F165 in EcNusG corresponds to F232 in AaNusG (Figure 1).

NusG is essential for bacterial growth (2, 13), on the basis of which the activity of nusG mutants was analyzed in a genetic assay (M. Bubunencko and D. L. Court, unpublished results). Nonessential genes can readily be replaced with an antibiotic-resistant gene using recombineering in *E. coli* (36). Similarly, when the same antibiotic-resistant cassette is used to replace an essential gene, drug-resistant recombinants occur but at a much (1000-fold) lower frequency. These recombinants contain a duplication of the essential gene with one allele being replaced by the drug-resistant cassette and the other being normal. When the same replacement is carried out in the presence of a plasmid carrying the wild-type gene, a simple replacement of the single chromosomal gene can

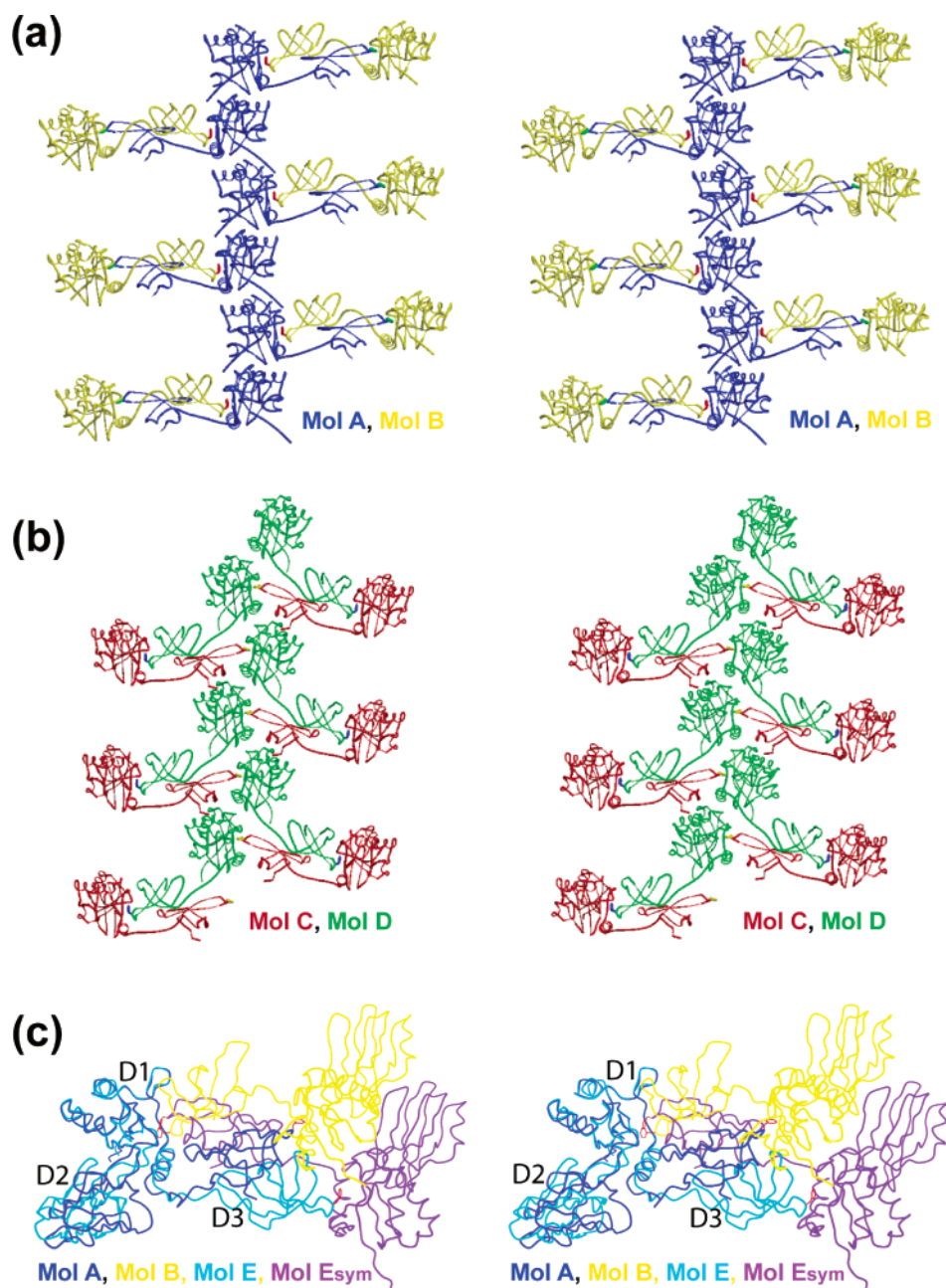


FIGURE 4: Stereoviews showing the oligomeric states of AaNusG with F232 (ball) highlighted. (a) Mol A (blue) and B (yellow) form a domain-swapped dimer, six of which are shown. (b) Mol C (red) and D (green) form the repeating unit of a polymer. A polymer with six repeating units is shown. (c) Mol E (cyan) forms a domain-swapped dimer with a symmetry-related molecule, E_{sym} (purple), which is different from the Mol A–Mol B dimer. The alignment of the two dimers is based on the optimized superposition of D1. This figure was prepared with Molscript (39) and Raster3D (40).

occur. If the plasmid gene carries a lethal mutation and therefore is not able to functionally replace the chromosomally disrupted gene, only duplications will be selected. If it is not lethal and fully complementary for function, the chromosomal copy can be replaced. Thus, only the haploid disrupted gene for viable mutants and the duplicated configuration for lethal mutants will be selected. These two forms of drug-resistant recombinants can be readily analyzed by PCR amplification of the chromosomal *nusG* gene (37).

RESULTS AND DISCUSSION

Overall Structure. Four independent molecules (Mol A–D) and one molecule (Mol E) are present in the asymmetric unit of the $P2_1$ and $C222_1$ crystals, respectively.

Each molecule is comprised of three domains, D1–D3 (Figure 2a). D1 contains three α -helices, four β -strands, and one 3_{10} -helix and forms a four-stranded antiparallel β -sheet flanked by one α -helix on one side of the β -sheet and two α -helices on the other. D1 and D2 are connected through $\beta 3$ and $\beta 10$ (Figure 2a). D2 consists of seven β -strands and two 3_{10} -helices and is organized into two sandwiched antiparallel β -sheets. D3 contains five β -strands and two 3_{10} -helices, forming a highly bent β -sheet. D1 and D3 are loosely tethered by a flexible six-residue linker (residues 187–192) (Figure 2a).

The AaNusG molecule is intrinsically flexible. Four to nine residues are disordered at the N-termini of the five molecules, and one or two residues are missing at the C-termini of Mol

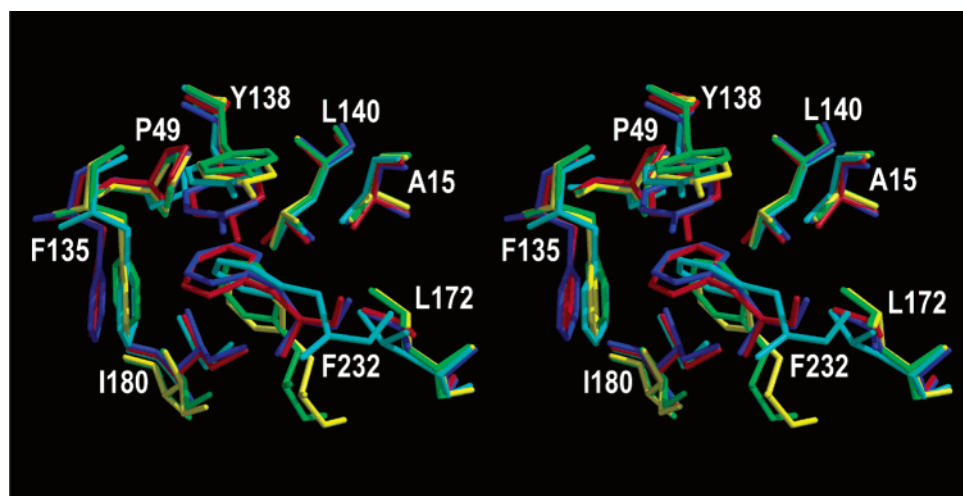


FIGURE 5: Stereoview of the five ball-and-socket junctions. F232 in D3 is the ball; A15, P49, F135, Y138, L140, L172, and I180 in D1 form the socket. $D1_{\text{Mol A}}-D3_{\text{Mol B}}$, $D1_{\text{Mol B}}-D3_{\text{Mol A}}$, $D1_{\text{Mol C}}-D3_{\text{Mol D}}$, $D1_{\text{Mol D}}-D3_{\text{Mol C}_{\text{sym}}}$, and $D1_{\text{Mol E}}-D3_{\text{Mol E}_{\text{sym}}}$ junctions are shown in blue, yellow, red, green, and cyan, respectively. This figure was prepared with Molscrip (39) and Raster3D (40).

C and Mol D. Three of five six-residue linkers can be located in the initial experimental map from MAD phasing; all five are defined unambiguously. Although every individual domain of the five molecules aligns quite well, the entire molecules do not. Figure 3 depicts the alignment of the five independent molecules based on the optimized superposition of $C\alpha$ atoms in D1. Among the five molecules, significant dislocations are observed for D2, while dramatic dispositions are found for D3. The $\beta 3-\beta 10$ connection between D1 and D2 allows the rotation of D2 around their long axes by $\sim 15^\circ$, whereas the six-residue linker allows the rotation of D3 along their long axes by $\sim 180^\circ$. Furthermore, the partial unwinding of the last turn of $\alpha 3$ in D1 allows D3 in Mol D to point in the opposite direction of its counterparts in Mol A–C and E (Figure 3). The conformation of Mol D in our $P2_1$ crystal is distinctly different from that in the previously reported $P2_1$ crystal (21), where the alignment of the four molecules shows that D3 of Mol D points in the same direction as others (Figure 2E in ref 21).

Oligomeric Forms and the Ball-and-Socket Junction. In the $P2_1$ lattice, Mol A and Mol B form a domain-swapped dimer (Figure 4a), whereas Mol C and Mol D compose the repeating unit of a polymer (Figure 4b). In the $C222_1$ lattice, Mol E forms a domain-swapped dimer with a symmetry-related molecule, Mol E_{sym} (Figure 4c). Note that the Mol A–Mol B dimer is different from the Mol E–Mol E_{sym} dimer. In $P2_1$ crystal, Mol A and Mol B are different due to the rotation of D3 along their long axes by $\sim 100^\circ$ (Figure 3), and the resulting dimer is not symmetric; whereas in the $C222_1$ crystal, Mol E and Mol E_{sym} are identical, and the resulting dimer is symmetric (Figure 4c). Moreover, the Mol E–Mol E_{sym} dimer is flat, and the Mol A–Mol B dimer is bent (Figure 4c). Perhaps further bending of the domain-swapped dimer results in the Mol C–Mol D pair, which is the repeating unit of a polymer (Figure 4b). Despite the different oligomeric states, the intermolecular interactions are the same. A ball-and-socket junction defines both types of dimers as well as the polymer. The ball is the side chain of F232, which is located at the tip of D3 in the loop region between $\beta 15$ and $\beta 16$ (Figure 2a). The side chains of A15, P49, F135, Y138, L140, L172, and I180 form the hydrophobic socket, which is located on D1 and between $\alpha 3$ and

$\beta 12$ (Figure 2). Note that the residues that form the ball and the socket are highly conserved in the *A. aeolicus* and *E. coli* proteins (Figure 1). Mol A and B form the domain-swapped dimer via two junctions, one of which is between $D3_{\text{Mol A}}$ and $D1_{\text{Mol B}}$ and the other between $D3_{\text{Mol B}}$ and $D1_{\text{Mol A}}$ (Figure 4a). Mol E and E_{sym} form the domain-swapped dimer in a similar manner, although its conformation is significantly different from that of the Mol A–Mol B dimer (Figure 4c). Within the Mol C–Mol D repeating unit of the polymer, only one ball-and-socket junction exists between $D1_{\text{Mol C}}$ and $D3_{\text{Mol D}}$ (Figure 4b). $D1_{\text{Mol D}}$ does not interact with $D3_{\text{Mol C}}$; instead, it interacts with $D3_{\text{Mol C}_{\text{sym}}}$ (Figure 4b). Strikingly, the domain-swapped dimers, the repeating unit of the polymer, and the polymer are all formed by the same unique ball-and-socket junction between D1 and D3 (Figure 4).

The formation of the intermolecular ball-and-socket junction between D1 and D3 buries a total surface area of $\sim 1000 \text{ \AA}^2$. Figure 5 depicts the alignment of five distinct ball-and-socket junctions observed in the two crystals, including $D1_{\text{Mol A}}-D3_{\text{Mol B}}$ and $D1_{\text{Mol B}}-D3_{\text{Mol A}}$ junctions (Figure 4a), $D1_{\text{Mol C}}-D3_{\text{Mol D}}$ and $D1_{\text{Mol D}}-D3_{\text{Mol C}_{\text{sym}}}$ junctions (Figure 4b), and a $D1_{\text{Mol E}}-D3_{\text{Mol E}_{\text{sym}}}$ junction (Figure 4c). Note that the $D1_{\text{Mol E}_{\text{sym}}}-D3_{\text{Mol E}}$ junction is the same as the $D1_{\text{Mol E}}-D3_{\text{Mol E}_{\text{sym}}}$ junction. Within the socket, the phenyl ring of F232 (the ball) not only shifts its center by $\sim 1.5 \text{ \AA}$ but also rotates by $\sim 90^\circ$, mimicking indeed the movement of the ball in the socket of a mechanical ball-and-socket junction (Figure 5). Interestingly, the socket also shows significant binding-induced fit (Figure 5). No conserved hydrogen bonds are found between the ball side and the socket side residues (not shown). Therefore, the conserved ball-and-socket junction is the only dominant interaction between D1 and D3. We believe that this junction is a clue to the function of AaNusG.

Role of F232. Residue F232 of AaNusG corresponds to F165 of EcNusG (Figure 1). To assess the importance of the ball-and-socket interaction, five site-directed mutants at this position in EcNusG were constructed: F165A, F165D, F165T, F165V, and F165Y. The mutants were tested for NusG activity in supporting the growth of *E. coli*. The conservative replacements of F165 with V or Y fully

supported cell growth, whereas a replacement with D or T made NusG defective for growth. The F165A mutant exhibited impaired growth at 37 °C and displayed a cold-sensitive phenotype at temperatures below 30 °C. This is the first conditional mutant reported for NusG. We conclude that the integrity of the ball-and-socket interaction is strongly dependent on the hydrophobicity and size of the ball element, which is vital for the growth of *E. coli*. We have recently identified a structure with similar properties in RNase III protein and carried out similar mutagenesis studies for the ball element (38). It appears that the hydrophobic character and size of the ball element are common determining factors of the ball-and-socket structures in these two proteins.

AaNusG as a Spring-Loaded Probe in Solution. In the crystal, AaNusG is an extended nonglobular molecule (Figure 2a) that is stabilized via dimerization and polymerization (Figure 4). If the extended and nonglobular form of the molecule were present in solution, the hydrophobic surface (~1000 Å²) buried upon the formation of the ball-and-socket junction would be exposed to solvent and the retention time of such molecules should be significantly different from that of globular molecules. However, our size exclusion chromatography experiments indicate that the average retention volume for AaNusG (MW = 28 kDa) is between those for the standards RNase A (MW = 13.7 kDa) and ovalbumin (MW = 44 kDa), suggesting that the AaNusG molecule adopts a more globular form in solution.

We propose that the long flexible six-residue linker (residues 187–192) and the ball-and-socket junction are two of the key factors for the structure and function of the molecule. Molecular modeling suggests that the ball in D3 can easily reach the socket in D1 to form a ball-and-socket junction within the same molecule (Figure 2b), mimicking a spring-loaded probe. The relaxed structure observed in the crystal would correspond to the sprung conformation, with which NusG interacts with other biological molecules. The fact that the spring-loaded conformation can be disrupted by the weak forces involved in crystallization suggests that the spring-loaded probe is unstable and readily releases the third domain to interact with other biomolecules, the nature of which remains to be elucidated.

ACKNOWLEDGMENT

We thank Dr. Z. Dauter for assistance during X-ray data collection, Dr. A. Evdokimov for help with the expression of SeMet protein, and Drs. A. Wlodawer and M. Kashlev for insightful discussions and critical reading of the manuscript.

REFERENCES

- Das, A., Pal, M., Mena, J. G., Whalen, W., Wolska, K., Crossley, R., Rees, W., von Hippel, P. H., Costantino, N., Court, D., Mazzulla, M., Altieri, A. S., Byrd, R. A., Chattopadhyay, S., DeVito, J., and Ghosh, B. (1996) *Methods Enzymol.* 274, 374–402.
- Zhou, Y., Filter, J. J., Court, D. L., Gottesman, M. E., and Friedman, D. I. (2002) *J. Bacteriol.* 184, 3416–3418.
- Li, J., Horwitz, R., McCracken, S., and Greenblatt, J. (1992) *J. Biol. Chem.* 267, 6012–6019.
- Sullivan, S. L., and Gottesman, M. E. (1992) *Cell* 68, 989–994.
- Li, J., Mason, S. W., and Greenblatt, J. (1993) *Genes Dev.* 7, 161–172.
- Nehrke, K. W., and Platt, T. (1994) *J. Mol. Biol.* 243, 830–839.
- Pasman, Z., and von Hippel, P. H. (2000) *Biochemistry* 39, 5573–5585.
- Burns, C. M., Nowatzke, W. L., and Richardson, J. P. (1999) *J. Biol. Chem.* 274, 5245–5251.
- Burns, C. M., Richardson, L. V., and Richardson, J. P. (1998) *J. Mol. Biol.* 278, 307–316.
- Burns, C. M., and Richardson, J. P. (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92, 4738–4742.
- Burova, E., and Gottesman, M. E. (1995) *Mol. Microbiol.* 17, 633–641.
- Burova, E., Hung, S. C., Sagitov, V., Stitt, B. L., and Gottesman, M. E. (1995) *J. Bacteriol.* 177, 1388–1392.
- Burova, E., Hung, S. C., Chen, J., Court, D. L., Zhou, J. G., Mogilnitskiy, G., and Gottesman, M. E. (1999) *Mol. Microbiol.* 31, 1783–1793.
- Downing, W. L., Sullivan, S. L., Gottesman, M. E., and Dennis, P. P. (1990) *J. Bacteriol.* 172, 1621–1627.
- Sullivan, S. L., Ward, D. F., and Gottesman, M. E. (1992) *J. Bacteriol.* 174, 1339–1344.
- Kypides, N. C., Woese, C. R., and Ouzounis, C. A. (1996) *Trends Biochem. Sci.* 21, 425–426.
- Zellers, M., and Squires, C. L. (1999) *Mol. Microbiol.* 32, 1296–1304.
- Squires, C. L., and Zaporozets, D. (2000) *Annu. Rev. Microbiol.* 54, 775–798.
- Hartzog, G. A., Wada, T., Handa, H., and Winston, F. (1998) *Genes Dev.* 12, 357–369.
- Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G. A., Winston, F., Buratowski, S., and Handa, H. (1998) *Genes Dev.* 12, 343–356.
- Steiner, T., Kaiser, J. T., Marinkovic, S., Huber, R., and Wahl, M. C. (2002) *EMBO J.* 21, 4641–4653.
- Andrykovitch, M., Guo, W., Routzahn, K. M., Gu, Y., Anderson, D. E., Reshetnikova, L. S., Knowlton, J. R., Waugh, D. S., and Ji, X. (2002) *Acta Crystallogr. D* 58, 2157–2158.
- Double, S. (1997) *Methods Enzymol.* 276, 523–530.
- Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* 276, 307–326.
- Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. D* 55, 849–861.
- de La Fortelle, E., and Bricogne, G. (1997) *Methods Enzymol.* 276, 472–494.
- Cowtan, K. (1994) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* 31, 34–38.
- Abrahams, J. P., and Leslie, A. G. W. (1996) *Acta Crystallogr. D* 52, 30–42.
- Lamzin, V. S., and Wilson, K. S. (1997) *Methods Enzymol.* 277, 269–305.
- 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. D* 54, 905–921.
- Jones, T. A., and Kjeldgaard, M. (1997) *Methods Enzymol.* 277, 173–208.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291.
- Navaza, J. (1994) *Acta Crystallogr. A* 50, 157–163.
- Bubeck, P., Winkler, M., and Bartsch, W. (1993) *Nucleic Acids Res.* 21, 3601–3602.
- Zhang, Y., Buchholz, F., Muirers, J. P., and Stewart, A. F. (1998) *Nat. Genet.* 20, 123–128.
- Court, D. L., Sawitzke, J. A., and Thomason, L. C. (2002) *Annu. Rev. Genet.* 36, 361–388.
- Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G., and Court, D. L. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97, 5978–5983.
- Błaszczak, J., Tropea, J. E., Bubunenkov, M., Routzahn, K. M., Waugh, D. S., Court, D. L., and Ji, X. (2001) *Structure* 9, 1225–1236.
- Kraulis, P. J. (1991) *J. Appl. Crystallogr.* 24, 946–950.
- Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* 277, 505–524.