NSF N-Terminal Domain Crystal Structure: Models of NSF Function

Richard C. Yu,* Reinhard J ahn,¹ and Axel T. Brunger*²
*The Howard Hughes Medical Institute and Department of Molecular Biophysics
and Biochemistry
Yale University
New Haven, Connecticut 06520
¹Department of Neurobiology
Max-Planck-Institute for Biophysical Chemistry
D-37077 Göttingen
Germany

Summary

N-ethylmaleimide-sensitive factor (NSF) is a hexameric ATPase essential for eukaryotic vesicle fusion. Along with SNAP proteins, it disassembles cis-SNARE complexes upon ATP hydrolysis, preparing SNAREs for trans complex formation. We have determined the crystal structure of the N-terminal domain of NSF (N) to 1.9 Å resolution. N contains two subdomains which form a groove that is a likely SNAP interaction site. Unexpectedly, both N subdomains are structurally similar to domains in EF-Tu. Based on this similarity, we propose a model for a large conformational change in NSF that drives SNARE complex disassembly.

Introduction

N-ethylmaleimide-sensitive factor (NSF), a member of the ATPases associated with diverse cellular activities (AAA) superfamily (Patel and Latterich, 1993; Confalonieri and Duguet, 1995), is involved in many eukaryotic membrane fusion pathways, from intra-Golgi transport to synaptic vesicle exocytosis (Bennett and Scheller, 1993; Ferro-Novick and J ahn, 1994; Rothman, 1996). These fusion events require SNAREs (SNAP receptors), membrane proteins whose cytosolic portions associate to form stable complexes, promoting and possibly driving the fusion event (Hay and Scheller, 1997; Weber et al., 1998). NSF, in an ATP hydrolysis-dependent manner, primes membranes for fusion by disrupting cis complexes of SNAREs, freeing them for recycling and activating them for formation of intermembrane trans complexes. These disassembly events require the cofactor soluble NSF attachment protein (SNAP) to bind to the SNAREs, providing a binding site for NSF (Whiteheart et al., 1992; Wilson et al., 1992). SNAP/SNARE binding stimulates NSF’s ATP hydrolysis rate, and this stimulation is coincident with complex disassembly (Morgan et al., 1994; Barnard et al., 1997). In detergent-solubilized cell and tissue extracts (Wilson et al., 1992; Hayashi et al., 1994) and in vitro (Söllner et al., 1993; Hayashi et al., 1995), a distinct complex of SNAREs, α-SNAP (a specific isoform of the SNAP family), and NSF can be isolated in the presence of nonhydrolyzable ATP analogs. This 205 particle, named for its sedimentation characteristics, likely represents an intermediate in the NSF-mediated SNARE complex disassembly pathway.

NSF is a hexamer (Fleming et al., 1998; Lenzen et al., 1998; Yu et al., 1998), and only the oligomeric form is competent for SNARE complex disassembly (Whiteheart et al., 1994; Nagiec et al., 1995). Each NSF protomer consists of three domains, defined by limited proteolysis: N (residues 1–205), D1 (206–477), and D2 (478–744) (Tayaga et al., 1993). Domain deletion and swapping show that N is necessary, but not sufficient, for the binding and disassembly activities of NSF (Nagiec et al., 1995). D1 and D2 each contain an AAA cassette, a consensus sequence of around 230 amino acids characteristic of AAA proteins (Confalonieri and Duguet, 1995). This cassette includes common nucleotide-binding motifs, such as the Walker A and B boxes (Walker et al., 1982), and additional characteristic regions (Beyer, 1997). ATP binding and hydrolysis by D1 are necessary for the SNARE disassembly reaction to occur, and ATP binding, but not hydrolysis, by D2 is necessary for hexamer formation (Sumida et al., 1994; Whiteheart et al., 1994).

In electron microscopy (EM) studies, the NSF hexamer in the presence of ADP appears as a stack of two rings, corresponding to the D1 and D2 domains, respectively, which is ~10 nm high and ~13 nm wide with an ~2.5 nm diameter pore along the symmetry axis (Hanson et al., 1997; Hohl et al., 1998). The shape of the hexamer depends on the nucleotide bound to D1 (Figure 7B). In the presence of ATP-EDTA or nonhydrolyzable ATP analog, the ~10 nm long × ~13 nm wide D1-D2 rings show diffuse additional features, probably N domains, flared outward away from the pore (Hanson et al., 1997). In this ATP-bound conformation, NSF can bind the SNAP/SNARE complex. The ADP-bound morphology is slightly wider, with the D1 ring dilated to ~16 nm (Hanson et al., 1997), and does not possess outwardly flared N domains.

A coherent picture of the NSF/SNAP/SNARE interactions is beginning to emerge based on crystallographic, biochemical, and EM studies. Crystal structures of a yeast homolog of α-SNAP (Rice and Brunger, 1999 [this issue of Molecular Cell]) and of the core of the synaptic SNARE complex (Sutton et al., 1998) have been determined. α-SNAP consists of an N-terminal twisted sheet of helix-loop-helix repeats, followed by a globular, α-helical bundle, and the core SNARE complex is a long, all-parallel four-helix bundle. Three α-SNAPs bind to the SNARE complex (Hayashi et al., 1995) in a lateral fashion, with the N terminus of each α-SNAP close to the membrane-proximal C termini of the SNAREs (Hohl et al., 1998). NSF then binds to the SNAP/SNARE complex at the C-terminal region of α-SNAP (Barnard et al., 1996). This region of α-SNAP includes a conserved leucine residue, Leu-294, which is crucial for both stimulation of NSF ATPase activity and efficient SNARE complex disassembly (Barnard et al., 1997). In the 205 complex,
Table 1. Proteins Structurally Similar to N

| Protein                          | Double- || | C, Rmsd (Å) |
|---------------------------------|-------|---|--------------|
| DMSO reductase                  | Y     | 5.8 | 2.6          |
| 1-aspartate-alpha-decarboxylase  | Y     | 5.7 | 2.3          |
| Formate dehydrogenase           | Y     | 5.6 | 2.3          |
| Barwin, basic barley seed protein | Y   | 4.0 | 2.8          |
| 1-fucose isomerase              | Y     | 3.9 | 2.9          |
| ELAV protease                   | Y     | 3.2 | 2.2          |
| Phthalate dioxygenase           | N     | 3.0 | 3.8          |
| Flavodoxin reductase            | N     | 3.0 | 4.1          |
| Elongation factor Tu            | N     | 3.0 | 2.9          |

N and DALI Structural Similarity Search Hits (All)

<table>
<thead>
<tr>
<th>Protein</th>
<th>Z</th>
<th>C, Rmsd (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elongation factor tu</td>
<td>2.9</td>
<td>3.2</td>
</tr>
</tbody>
</table>

NSF/E-Tu Alignments*

<table>
<thead>
<tr>
<th>Nsub</th>
<th>EF-Tu* Domain 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met-1 to Cys-11</td>
<td>Pro-220 to Thr-230</td>
</tr>
<tr>
<td>Thr-13 to Ser-17</td>
<td>Ile-231 to Gly-235</td>
</tr>
<tr>
<td>Asp-28 to Ser-54</td>
<td>Lys-246 to Met-272</td>
</tr>
<tr>
<td>Val-56 to Ser-59</td>
<td>Asp-284 to Gly-287</td>
</tr>
<tr>
<td>Gin-66 to Trp-69</td>
<td>Ser-294 to Glu-297</td>
</tr>
<tr>
<td>Gin-76 to Ser-84</td>
<td>Arg-300 to Gly-308</td>
</tr>
</tbody>
</table>

**Aligned strands and C, rmsd values determined by DALI (Holm and Sander, 1993). Z is defined as the number of standard deviations above expected for similarity between the aligned protein strands.**

The N-terminal domain has two subdomains, a double β barrel (Castillo et al., 1999), and a four strand-one helix α/β roll. The solvent-accessible surface features a large, very positively charged groove that could bind to the highly negative, α-helical, C-terminal region of α-SNAP (Rice and Brunger, 1999). Despite limited primary sequence similarity, both subdomains are structurally similar to adjacent domains in EF-Tu, a protein that exhibits large nucleotide hydrolysis-dependent conformational changes (Jumak and Abel, 1996). NSF and EF-Tu each contain a nucleotide-binding domain, D1 in NSF and domain 1 in EF-Tu, adjacent to these structurally similar subdomains. These similarities imply that both proteins may share a common hydrolysis-triggered conformational change. The insights obtained from this system will likely be generally applicable to other members of the AAA superfamily.

**Results and Discussion**

**Structure Determination**

Recombinant selenomethionine (Se-Met)-derivatized Chinese hamster ovary N was expressed in E. coli, purified to homogeneity, and crystallized by vapor diffusion. The crystal structure was solved to 1.9 Å resolution using multienzyme anomalous diffraction (MAD) data collected at four wavelengths (Hendrickson, 1991) around the selenium absorption edge. The three ordered Se-Met sites were found by an automated Patterson search method (Grosse-Kunstleve and Brunger, 1999) and histogram matching (Zhang and Main, 1990), an excellent, almost completely traceable, experimental electron density map was obtained (Figures 2A and 2B). The final model refined to an Rfree value of 24.2% and includes residues 1-161 and 167-201, three sulfate ions, and 130 waters. Residues 162-166 and 202-205 were disordered.

The N-terminal domain has two subdomains sharing a hydrophilic interface.
hydrophilic interface between the two subdomains bur-
ies approximately 1500 Å² and contains seven well-
ordered water molecules (Figure 2C). The mean shape 
complementarity index (\(<Sc\>) of the buried surfaces
(Lawrence and Colman, 1993) is 0.567, well below 0.70,
the average value expected for normal protein-protein 
interfaces; the values of \(<Sc\>) range from 0 to 1, for
zero to perfect surface complementarity, respectively.
Upon including the seven ordered water molecules,
\(<Sc\>) increases to 0.68, indicating that the buried water 
molecules substantially compensate for the lack of
shape complementarity between the two subdomains.
Additionally, two small hydrophobic clusters are found
at the edges of the interface, in the region between 
strand \(\beta6\) and helix \(\alpha1^1\), and the area formed by helices 
\(\alpha1\) and \(\alpha2\), and strand \(\beta4^\prime\) (Figure 2C). While the hydrated
interface between the two NSF subdomains suggests
that these surfaces could be solvent exposed at little
energetic cost, there is evidence that suggests that the
two subdomains are stably packed together and act as
a unit. First, the thermal factors of residues are lowest
the average value expected for normal protein-protein 
interfaces; the values of \(<Sc\>) range from 0 to 1, for
zero to perfect surface complementarity, respectively.
Upon including the seven ordered water molecules,
\(<Sc\>) increases to 0.68, indicating that the buried water 
molecules substantially compensate for the lack of
shape complementarity between the two subdomains.
Additionally, two small hydrophobic clusters are found
at the edges of the interface, in the region between 
strand \(\beta6\) and helix \(\alpha1^1\), and the area formed by helices
Figure 3. Structural Similarity between N and EF-Tu
Stereo backbone alignments of subdomains NA and NB to EF-Tu domains 2 and 3, respectively. Thick segments are considered equivalent by DALI and were used for Cα alignment. (A) NA from blue (N terminus) to pink (C terminus), EF-Tu domain 2 from dark green to light green. (B) NB from pink to red, EF-Tu from light green to yellow. Equivalent strands from NA and EF-Tu domain 2 align with a Cα rmsd of 2.9 Å, and those from NB and EF-Tu domain 3 align with a Cα rmsd of 3.2 Å. Red dashed line indicates disordered residues 162-166.

existing as a compact, single domain. A more complicated biphasic melting behavior would be expected if the structurally distinct NA and NB domains were separable in solution and denatured independently.

The N-Terminal Subdomain Is a Double-β Barrel
Structural similarity searches using DALI (Holm and Sander, 1993) with NA as a target yielded several proteins containing a recently described double-β barrel (DPBB) motif (Castillo et al., 1999), such as DMSO reductase, aspartate decarboxylase, formate dehydrogenase, barwin, and fucose isomerase (Table 1). In fact, NSF homologs such as CDC-48p and the transitional endoplasmic reticulum ATPase (TERA) were predicted, on the basis of primary sequence analyses, to contain a DPBB motif. The aligned strands between NA and the similar DPBB proteins overlap with Cα rmsd’s ranging from 2.3 to 2.9 Å. The DPBB motif contains two interlocking β loops, which are strand-loop arrangements roughly resembling the Greek letter β (Figure 1C). The first β loop consists of strand β5 and the loop between strands β1 and β2; the second consists of strand β2 and the loop between strands β4 and β5. The topology of NA is identical to that of the canonical DPBB except for a relocation of helix α1 from the consensus position between strands β2 and β3 to the first β loop region (Figure 1B). This disrupts the two-fold pseudosymmetry of the strand topology and places both helix α1 and α2
along one edge of a major surface groove in N. One common feature of DPBB motif-containing proteins is a catalytic or substrate-binding site in the $\psi$ loop regions, found between strands $\beta 1$ and $\beta 2$, and between strands $\beta 3$ and $\beta 4$ (Castillo et al., 1999). This may indicate a substrate-binding region in N at this $\psi$ loop.

Interestingly, domain 2 of EF-Tu is also structurally similar to N (Table 1; Figure 3B). Both are small six-stranded $\beta$ barrels with shear numbers of 10, though EF-Tu domain 2 is topologically a Greek key $\beta$ barrel rather than a DPBB. The core $\beta$ strands of the barrels in the EF-Tu domain 2 and in N$_{\gamma}$ align with a $C_\alpha$ root-mean-square difference (rmsd) of 2.9 Å (Table 1).

The C-Terminal Subdomain Is an $\alpha/\beta$ Roll N$_{\gamma}$ is a roll consisting of one $\alpha$ helix, $\alpha 3'$, and four $\beta$ strands, $\beta 1'$, $\beta 4'$, $\beta 5'$, and $\beta 7'$ (Figure 1A). Helix $\alpha 3'$ is capped by two small two-stranded $\beta$ sheets, antiparallel
Figure 5. Sequence Alignment of CHO NSF to Other Known Orthologs

CHO NSF aligned with mouse, human, Drosophila (two isoforms), C. elegans, yeast S. cerevisiae, and yeast C. albicans. Sequences were aligned using ClustalW (Thompson et al., 1994), then conservation values were calculated with AMAS (Livingstone and Barton, 1993). Colors indicated weak (light blue) to absolute (violet with white lettering) conservation. The top line indicates regions of β sheet strand (arrow) and α- or 310-helix (box) as assigned by DSSP (Kabsch and Sander, 1983). Numbers indicate residue number of CHO NSF. The dashed line indicates disordered regions.

Strands β3′ and β6′, and parallel strands β2′ and β8′. The core of the roll between helix α3′ and strands β1′, β4′, and β5′ is made up of a peculiar row of highly conserved phenylalanine and tyrosine side chains (Figure 2), which form a ladder of aromatic side chains (Figure 1D).

Surprisingly, another domain of EF-Tu, domain 3, was the sole structure in the representative structure database identified by DALI as being similar to N0 (Table 1). The topologies between N0 and domain 3 of EF-Tu, a six-stranded jelly roll β barrel, are not identical, but four β strands, consisting of 61 residues, closely superimpose with a Cα rms deviation of 3.2 Å (Figure 3; Table 1).

Surface Charge, Conservation, and Shape

The 10,500 Å² solvent-accessible surface is highly textured, with many knobs and grooves. It possesses an overall positive charge (Figure 4A) and, compared to known orthologs (Figure 5), shows concentrated patches of primary sequence conservation (Figure 4B). Positive charges are primarily localized to two areas: near the disordered loop in N0 between strands β5′ and β6′, and in the large groove formed by helices α1 and α2, and strand β4′ (Figure 4A). Negative charges are found on faces of the protein away from the large positive groove. Primary sequence conservation of the accessible surface clusters into two regions: a small patch near the disordered loop, and a much larger streak along the intersubdomain cavity, starting at the lower half of the groove along N0 and terminating on the other side of the protein along the surface of Nα.

A cavity analysis performed by SURFNET (Laskowski, 1995) using a 2.0-5.0 Å radius probe identified three grooves, all large enough to potentially accommodate a typical α helix (Figure 4C). Grooves 1 and 2 are above

Figure 6. Model of α-SNAP Binding to Groove 3 of N

Assuming that the position of D1 is near the C terminus of N, binding of the C terminus of α-SNAP to groove 3 would allow for the penultimate residue of α-SNAP to interact with D1.
Crystal Structure of the N-Terminal Domain of NSF

Figure 7. Model of Conformational Changes in NSF Based on the Structural Similarity to EF-Tu

(A) Nucleotide-dependent conformation of EF-Tu. The switch II helix, which changes conformation upon loss of the gamma phosphate of the bound nucleotide, is shown in purple. The nucleotide, drawn in ball-and-stick, is cradled by the conserved P loop region, drawn in red. The GMPPNP- (top) and GDP-bound (bottom) conformations are aligned by a least-squares fit of the P loop and nucleotide atoms in domain 1 (orange). Upon nucleotide hydrolysis, domains 2 (green) and 3 (yellow) twist and move back into the plane of the figure and remain fixed relative to each other during this rotation. During this movement, domain 1 swings approximately 90° about an axis through domains 2 and 3. PDB accession codes for GMPPNP- and GDP-bound EF-Tu are 1EFT (Kjeldgaard et al., 1993) and 1TUI (Polekhina et al., 1996), respectively.

(B) Model of nucleotide-dependent motions in an NSF hexamer, held together by a ring of D2 domains (red). In the ATP-charged state, the N domains (yellow and green) are flared outward from the pore of the hexamer. After ATP hydrolysis, N rotates and twists relative to D1 (orange). This model is consistent with estimated hexamer dimensions derived from EM studies (Hanson et al., 1997; Hohl et al., 1998).

The Putative α-SNAP Binding Site

The crystal structure of N and the recently solved structure of the yeast SNAP Sec17 (Rice and Brunger, 1999) suggest possible modes of interactions between NSF and α-SNAP. Assuming structural similarity between Sec17 and α-SNAP, the C-terminal portion of α-SNAP, necessary for NSF binding and ATPase stimulation (Barnard et al., 1996, 1997), is predominantly α-helical and negatively charged. Surface charge, sequence conservation, and shape limit potential α-SNAP-binding sites on N to two of the three surface cavities (Figure 4C). While all three are large enough to bind an α-helix, the predominantly negative charge of groove 2 (Figure 4A) makes it an unlikely binding site. Though groove 1 is weakly positively charged and conserved, groove 3 is more likely to interact with α-SNAP since it has the highest local concentration of positive charge and sequence conservation (Figure 4C). Interestingly, groove 3 contains the first N<sub>α</sub> loop, which serves as a binding or catalysis site in many DPBB proteins (Castillo et al., 1998).

Leucine 294 of α-SNAP, a residue implicated in D1 nucleotide hydrolysis stimulation (Barnard et al., 1997), is C-terminal to the ultimate α helix, just outside the region of ordered electron density in the crystal structure of Sec17 (Rice and Brunger, 1999). The hydrolysis-stimulating effect of this residue suggests a direct interaction into a slightly negatively charged concave region defined by helix α2 and strands β1', β4', and β5'.

A Model of NSF Domain Motion

The conformational changes observed by EM in NSF upon ATP hydrolysis may provide a mechanical force used to disassemble the SNARE complex. In light of the
### Table 2. Crystallographic Data and Phasing

<table>
<thead>
<tr>
<th>Crystal</th>
<th>Space Group</th>
<th>Cell Dimensions</th>
<th>Data Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Se-Met</td>
<td>C222</td>
<td>a = 32.04 Å, b = 175.36 Å, c = 79.66 Å, α = β = γ = 90°</td>
<td></td>
</tr>
<tr>
<td>Se-Met λ₁ (0.9799 Å)</td>
<td>1.9</td>
<td>32,855</td>
<td>3.6 (2.6)</td>
</tr>
<tr>
<td>Se-Met λ₁ (0.9801 Å)</td>
<td>1.9</td>
<td>33,269</td>
<td>3.9 (2.6)</td>
</tr>
<tr>
<td>Se-Met λ₁ (0.9998 Å)</td>
<td>1.9</td>
<td>33,216</td>
<td>3.2 (2.7)</td>
</tr>
<tr>
<td>Se-Met λ₁ (0.9537 Å)</td>
<td>1.9</td>
<td>33,652</td>
<td>3.6 (3.1)</td>
</tr>
</tbody>
</table>

### Figures of Merit (FOM)

<table>
<thead>
<tr>
<th>FOM</th>
<th>Two-wavelength (λ₁λ₂) MAD</th>
<th>λ₁ SAD</th>
<th>λ₂ SAD</th>
<th>λ₃ SAD</th>
<th>Combined FOM*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.36 (0.15)</td>
<td>0.17 (0.03)</td>
<td>0.04 (0.00)</td>
<td>0.05 (0.00)</td>
<td>0.39 (0.13)</td>
</tr>
</tbody>
</table>

* Values in parentheses are for the high-resolution bin, 1.97-1.90 Å.

### Two-Wavelength MAD Phasing Power

<table>
<thead>
<tr>
<th>Phasing Power</th>
<th>16.66-1.90 (1.98-1.90) Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>λ₁ → λ₂</td>
<td>λ₂ → λ₃</td>
</tr>
<tr>
<td>0.74 (0.17)</td>
<td>0.69 (0.29)</td>
</tr>
<tr>
<td>SAD Phasing Power*</td>
<td>16.66-1.90 (1.99-1.90) Å</td>
</tr>
<tr>
<td>λ₁ → λ₂</td>
<td>λ₂ → λ₃</td>
</tr>
<tr>
<td>1.00 (0.13)</td>
<td>0.39 (0.00)</td>
</tr>
</tbody>
</table>

* Figure of merit of phase probability distribution combined from one two-wavelength MAD phase distribution and three one-wavelength SAD phase distributions.
Crystallization, Data Collection, and Processing

Consisting of 100 mM Tris (pH 8.7), 2.0 M ammonium sulfate, and anomalous signal (and resuspended in 0.2 ml of mother liquor for use as a seed stock were refined against the MAD or SAD data. Phase probability distributions were then streak seeded using the seed stock. Crystals grew using solvent flipping (Abrahams and Leslie, 1996) and histogram parallelization. This is apparent in the observed diffraction ratios, where the largest dispersive difference occurred between the two remote wavelengths. Normally, this dispersive difference is expected to be rather small. This systematic error was presumably introduced by collecting the MAD data one wavelength at a time. Data were processed using DENZO (Otwinowski and Minor, 1997), and intensities were reduced and scaled using SCALEPACK (Otwinowski and Minor, 1997) (Table 1).

Structure Determination

Three of the five expected selenium sites of N were found in anomalous difference Patterson maps calculated at the peak anomalous (\(\lambda_s\)) wavelength using an automated Patterson heavy atom search method (Grosse-Kunstleve and Brünger, 1999). The two remaining sites, one at the N terminus and one in a disordered loop, were found using MAD-CAD (Advanced Light Source, Berkeley, CA) using a Quantum-4 CCD detector (Area Detector Systems Corporation). Inverse beam geometry was employed using 15° wedges. The diffraction data was collected one wavelength at a time. Data were processed using DENZO (Otwinowski and Minor, 1997), and intensities were reduced and scaled using SCALEPACK (Otwinowski and Minor, 1997) (Table 1).

Table 3. Refinement Statistics

<table>
<thead>
<tr>
<th>Resolution (Å)</th>
<th>Overall</th>
<th>50.0–1.9</th>
<th>50.0–4.09</th>
<th>3.25</th>
<th>2.84</th>
<th>2.58</th>
<th>2.39</th>
<th>2.25</th>
<th>2.14</th>
<th>2.05</th>
<th>1.97</th>
<th>1.90</th>
</tr>
</thead>
<tbody>
<tr>
<td>R value</td>
<td>21.0</td>
<td>22.5</td>
<td>19.4</td>
<td>20.4</td>
<td>21.8</td>
<td>21.3</td>
<td>20.5</td>
<td>19.9</td>
<td>20.0</td>
<td>21.3</td>
<td>24.0</td>
<td></td>
</tr>
<tr>
<td>Rfree value</td>
<td>24.2</td>
<td>24.0</td>
<td>21.8</td>
<td>26.8</td>
<td>24.2</td>
<td>25.9</td>
<td>24.0</td>
<td>24.5</td>
<td>23.2</td>
<td>24.5</td>
<td>28.1</td>
<td></td>
</tr>
<tr>
<td>Luzzati coordinate error</td>
<td>0.22 Å</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-validated Luzzati coordinate error</td>
<td>0.26 Å</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond length deviation</td>
<td>0.010 Å</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond angle deviation</td>
<td>1.42°</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improper angle deviation</td>
<td>0.79°</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dihedrals</td>
<td>25.34°</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average B factor</td>
<td>25.53 Å²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum B factor</td>
<td>10.72 Å²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum B factor</td>
<td>70.06 Å²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residues in core regions (%)</td>
<td>95.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residues in disallowed regions (%)</td>
<td>0.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of reflections</td>
<td>32,730</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(R = \frac{|F_{\text{obs}}| - |F_{\text{calc}}|}{|F_{\text{calc}}|}\) and \(R_{\text{free}}\) value is the R value obtained for a test set of reflections, consisting of a randomly selected 10% subset of the diffraction data, not used during refinement.

Crystal Structure of the N-Terminal Domain of NSF

TTAATATGCGGCGGCGGAGATGCAAGC and GAATTCCCTGATCA CTCGAGGCCATTTCTCTGGTTTAGC. N was subcloned into the pET15b expression vector (Novagen). The resulting protein contained a thrombin-cleavable N-terminal hexahistidine tag. All DNA sequences were verified by dyeoxynucleotide sequencing (HHMI Biopolymer and W. M. K. B. Technology Research Laboratory, Yale University). The expression construct was transformed into BL21 (DE3) (Novagen). Cells were grown in M9 minimal medium containing 150 mg/l ampicillin at 37°C. At an optical density of \(A_{600nm} = 2.0\), methionine synthesis was inhibited by adding 100 mg/l D-lysine, D-phenylalanine, and D-threonine; 50 mg/l D-isoleucine and D-valine; and 60 mg/l D/L selenomethionine (Sigma). After 15 min, cells were induced with 0.4 mM isopropyl-1-\(\beta\)-D-thiogalactopyranoside. Four hours later, cells were harvested by centrifugation and stored at ~80°C. Frozen cells (20 g) were thawed and homogenized in 100 ml of buffer A (20 mM HEPES [pH 7.8], 500 mM NaCl, 10 mM 2-mercaptoethanol) and 0.1% Triton X-100, then lysed by sonication. The lysate was centrifuged for 30 min at 40,000 × g. The supernatant was loaded by gravity over a column packed with Mono S column (Pharmacia) and eluted with a 10 column volume linear gradient from 0% buffer C to 100% buffer C + 300 mM NaCl. N eluted at approximately 150 mM NaCl. Fractions containing N were concentrated to 10 mg/ml and dialyzed against buffer C + 150 mM NaCl for 12 hr. Protein concentrations were determined using the Bradford protein concentration assay (Bio-Rad) calibrated with a bovine serum albumin standard.

Crystalization, Data Collection, and Processing

Initial crystals grew by hanging drop vapor diffusion at 20°C. Two microliters of 10 mg/ml protein was mixed with 2 μl of mother liquor, consisting of 100 mM Tris (pH 8.7), 2.0 M ammonium sulfate, and 10 mM DTT, on a siliconized glass cover slip and equilibrated against 0.5 ml of mother liquor. Rod-shaped crystals grew to full size (200 μm x 40 μm x 40 μm) in 48 hr. Crystals were cryoprotected by sequential transfer into mother liquor with 5%–30% glycerol, and then flash frozen in liquid nitrogen-cooled propane.

N crystallized in space group C222, with one molecule per asymmetric unit. The solvent content of the crystal was approximately 52%. MAD data were collected using four wavelengths (Table 2) at 100 K at beamline 5.0.2 at the Advanced Light Source (Berkeley, CA) using a Quantum-4 CCD detector (Area Detector Systems Corporation). Inverse beam geometry was employed using 15° wedges. The diffraction data was collected one wavelength at a time. Data were processed using DENZO (Otwinowski and Minor, 1997), and intensities were reduced and scaled using SCALEPACK (Otwinowski and Minor, 1997) (Table 1).
matching (Zhang and Main, 1980). Density modification proved essential for producing traceable electron density maps. This procedure dramatically improved map quality and compensated significantly for the lack of dispersive information. All heavy atom search, phasing, and density modification calculations were carried out with the crystallography and NMR System (CNS) (Brunger et al., 1998).

Model Building and Refinement
The initial model was built using the program O (Jones et al., 1991). The high quality of the initial experimental electron density map and known selenium sites allowed unambiguous tracing for most of the protein backbone and side chain atoms. Progress was monitored with the Rmea value using a 10% randomly selected test set. This test set was also used for cross-validation of r2 and other statistical quantities required for the calculation of the maximum likelihood target and r2-weighted electron density maps. Initial refinement consisted of one iteration of torsion angle dynamics simulated annealing (Rice and Brunger, 1994) using the MLHL target function (Pannu and Read, 1998) with the experimental, density-modified phases as a prior phase distribution, followed by model rebuilding in O. Later, refinement consisted of iterative rounds of model building (including selecting chemically reasonable water molecules and sulfate ions) in phase-combined r2-weighted 2Fobs − Fcalcd maps (Read, 1986), conjugate gradient minimization, and individual restrained atomic B factor refinement (Hendrickson, 1985). Refinement was carried out against low energy remote wavelength 1/A2) diffraction data. Refinement used a flat bulk solvent correction (electron density level, 1/A2 = 0.40 e/A3; B = 46.3 Å2) (Jiang and Brunger, 1994) and overall anisotropic B factor correction (Biso = 5.54 Å2, B = 3.62 Å2, B12 = −9.16 Å2, B23 = 0.00 Å2, B31 = 0.00 Å2, B13 = 0.00 Å2) at 50–1.9 Å resolution. Final model statistics are shown in Table 3. The final model contains 130 water molecules and three sulfate ions. All refinement calculations were carried out with CNS (Brunger et al., 1998).

Figure Preparation
All figures were prepared using GL_Render (courtesy of Dr. L. Esser), Bobscript (Essnouf, 1997), Molscript (Kaasrus, 1991), and Pov-Ray (The POV-ray Team, 1998) unless otherwise noted.

Acknowledgments
We thank R. B. Sutton for expert advice; L. Rice, C. Strope, M. Bowen, M. Wilson, J. Ernst, M. Reese, and J. Hyman for stimulating discussions and critical reading of the manuscript; and T. Earnest, L.-W. Hung, C. Strope, and L. Rice for help with data collection at beamline 5.0.2 at the Advanced Light Source (ALS is funded by the US Department of Energy Office of Basic Energy Sciences); and P. Adams and R. Grosse-Kunstleve for expert help with CNS.

Received April 29, 1999; revised June 7, 1999.

References


Crystal Structure of the N-Terminal Domain of NSF


Protein Data Bank Accession Number

The accession number for the structure reported in this paper is 1QCS.