Structural Characterization of Hsp12, the Heat Shock Protein from *Saccharomyces cerevisiae*, in Aqueous Solution Where It Is Intrinsically Disordered and in Detergent Micelles Where It Is Locally $\alpha$-Helical

Kiran K. Singarapu, Marco Tonelli, Darius C. Chow, Ronnie O. Frederick, William M. Westler, and John L. Markley

Received for publication, September 22, 2011. Published, JBC Papers in Press, October 13, 2011, DOI 10.1074/jbc.M111.306464

Structural characterization of Hsp12 by NMR and biophysical techniques. The protein was labeled uniformly with nitrogen-15 and carbon-13 so that its conformation could be determined in detail both in aqueous solution and in two membrane-mimetic environments, SDS and dodecylphosphocholine (DPC) micelles. Secondary structural elements determined from assigned chemical shifts indicated that Hsp12 is dynamically disordered in aqueous solution, whereas it gains four helical stretches in the presence of SDS micelles and a single helix in the presence of DPC. These conclusions were reinforced by circular dichroism spectra of the protein in all three environments. The lack of long range interactions in NOESY spectra indicated that the helices present in SDS micelles do not pack together. $R_1$ and $R_2$, relaxation and heteronuclear NOE measurements showed that the protein is disordered in aqueous solution but becomes more ordered in the presence of detergent micelles. NMR spectra collected in the presence of paramagnetic spin relaxation agents (SDSA, 16DSA, and Gd(DTPA-BMA)) indicated that the amphipathic $\alpha$-helices of Hsp12 in SDS micelles lie on the membrane surface. These observations are in agreement with studies suggesting that Hsp12 functions to protect the membrane from desiccation.

Heat shock protein 12 (Hsp12), a 109-residue protein (12 kDa) encoded by gene HSP12 in *Saccharomyces cerevisiae* (UniProtKB ID: P22943/HSP12_YEAST), belongs to Pfam family PF04119, which currently contains 74 members with unknown three-dimensional structure. Hsp12 is part of a group of small heat shock proteins (Hsp) that function as chaperone proteins and are ubiquitously involved in nascent protein folding, typically by protecting proteins from misfolding (1–5). The transcriptionally regulated gene expression of Hsp proteins increases when the cells are exposed to extreme environments, such as high and low temperature or other stress (6). Accordingly, Hsp12 is produced by cells in early stationary phase and can be massively induced by oxidative or osmotic stress or by exposing cells to low or high temperatures or to high concentrations of sugar or ethanol (7, 8). Hsp12 has been characterized as a late embryonic abundant-like protein on the basis of its high content of hydrophilic amino acids, high solubility at very high temperatures (80 °C), and induction by osmotic shock (9–11). It has been shown that the Hsp12 gene is activated by the high osmolarity glycerol pathways and negatively regulated by protein kinase A (PKA) (12). Mutational studies suggest that Hsp12 is essential to the formation of a biofilm by the Sardinian yeast strain of yeast (13), which contributes to the tolerance of this yeast strain to freezing and high alcohol concentrations (14).

Trehalose is a disaccharide carbohydrate known to play an important role in protecting the cell membrane from dehydration. Trehalose replaces bound water in the membrane and stabilizes the intracellular water structure to preserve the structure and functional integrity of the plasma membrane (15–18).
Accumulation of high levels of trehalose and Hsp12 induction in both stressed and stationary phase cells have been implicated in the protective response to heat, cold, high concentration of alcohol, or desiccation. Because both trehalose synthesis and Hsp12 induction are increased by exposing the cells to the same extreme conditions, it has been suggested that they function synergistically. In fact, deletion of the Hsp12 gene results in the accumulation of increased levels of trehalose to maintain intracellular hydration (19).

We report results from solution nuclear magnetic resonance (NMR) spectroscopy, paramagnetic spin relaxation experiments, and circular dichroism (CD) spectroscopy that provide detailed information on the conformational state of Hsp12 in aqueous solution and in a membrane-mimetic environment. In aqueous solution, Hsp12 is dynamically disordered, whereas it forms one helix in the presence of DPC micelles and four helices in the presence of SDS that lie on the surface of the micelle. Another group that carried out more qualitative NMR investigations of Hsp12 in aqueous solution and in the presence of SDS micelles also reported ordering of the protein in the membrane-type environment (20).

**EXPERIMENTAL PROCEDURES**

**Protein Production and Sample Preparation**—We expressed the Hsp12 gene from yeast and purified the protein product by the pipeline protocol developed by the Center for Eukaryotic Structural Genomics (21). Briefly, we produced protein samples containing a cleavable N-terminal His6 tag in *Escherichia coli* cells grown in media supplemented either with $^{15}\text{N}$H$_4$Cl, for uniform labeling with nitrogen-15, or with $^{15}\text{N}$H$_4$Cl and $^{13}\text{C}$-glucose, for uniform labeling with both nitrogen-15 and carbon-13. We isolated protein samples by two-step immobilized metal ion affinity chromatography and added tobacco etch virus protease to cleave the purification tag, which we separated from the Hsp12 protein by gel filtration chromatography. The molecular masses of the purified $^{15}\text{N}$-labeled and $^{15}\text{N},^{13}\text{C}$-labeled proteins were determined at the University of Wisconsin-Madison Biotechnology Center by electrospray ionization mass spectrometry on an Applied Biosystems 3200 Q Trap LC/MS/MS system and by matrix-assisted laser desorption/ionization on an Applied Biosystems 4800 MALDI TOF/TOF. Protein identification was carried out by tryptic proteolysis and molecular weight assignment of the generated peptides by using an Agilent 1100 series nanoLC/MSD Trap SL.

We used three paramagnetic spin labels, 5DSA (5-doxyl steric acid free radical), 16DSA (16-doxyl steric acid free radical), and Gd(DTPA-BMA), to identify the interaction between the detergent and the protein. 5DSA and 16DSA were purchased from Sigma-Aldrich, and Gd(DTPA-BMA) was purchased from GE Healthcare, which markets it under the trade name Omniscan™. Because of their poor aqueous solubility, we dissolved 5DSA and 16DSA separately in methanol and dried the samples overnight under vacuum. To each of these samples (10 mM final spin label concentration) we added 0.2 mM [U-$^{15}\text{N}]$Hsp12 in the presence of 200 mM SDS and incubated the solutions overnight at 40 °C. We added a solution of

**FIGURE 1.** Hsp12 is intrinsically disordered in aqueous solution but becomes partly structured in the presence of SDS micelles. A, assigned $^1\text{H},^{15}\text{N}$ HSQC spectrum of Hsp12 in aqueous solution at 25 °C. B, assigned $^1\text{H},^{15}\text{N}$ HSQC spectrum of Hsp12 in the presence of 200 mM SDS micelles at 40 °C. C, assigned $^1\text{H},^{15}\text{N}$ HSQC spectrum of Hsp12 in the presence of 200 mM DPC at 40 °C. All spectra were recorded on a 900 MHz Varian spectrometer equipped with a cryogenic probe.

**FIGURE 2.** CD spectrum of Hsp12 as a function of added SDS micelles.
Structural Characterization of Hsp12

Gd(DTPA-BMA) directly to yield a sample consisting of 10 mM spin label, 0.2 mM [U-15N]Hsp12, and 200 mM SDS.

**NMR Spectroscopy**—We recorded all NMR spectra at the National Magnetic Resonance Facility at Madison on Varian VNMRS (600 MHz and 900 MHz) spectrometers equipped with triple-resonance cryogenic probes. The temperature of the sample was regulated at 25 °C for Hsp12 in aqueous solution and 40 °C for Hsp12 in the presence of 200 mM SDS (Hsp12_SDS) or 200 mM DPC (Hsp12_DPC). We collected a series of two- and three-dimensional NMR spectra (22) at 600 MHz for Hsp12, Hsp12_SDS, and Hsp12_DPC for use in assignments. All three samples contained 0.7 mM [U-13C,15N]Hsp12 in NMR buffer with 10 mM MOPS, 100 mM NaCl, 3 mM NaN3, 5 mM DTT, pH 7.0, 95% H2O, 5% D2O; the Hsp12_SDS and Hsp12_DPC samples contained in addition, respectively, 200 mM SDS or 200 mM DPC. We collected 13C-resolved three-dimensional 1H-1H NOESY and 15N-resolved three-dimensional 1H-1H NOESY data on 900 MHz for Hsp12_SDS and 600 MHz for Hsp12_DPC for deriving distance constraints. Raw NMR data were processed with NMRPipe (23) and analyzed using the program XEASY (24). Two-dimensional 1H-13C HSQC and three-dimensional HC(CO)NH, and three-dimensional C(CO)NH experiments were used to assign the side chain and HB and HA resonances for Hsp12_SDS and Hsp12_DPC. Three-dimensional 15N-edited 1H-1H NOESY (100-ms mixing time), and three-dimensional 13C-edited 1H-1H NOESY (120 ms) experiments were used to derive the distance constraints to determine the three-dimensional structure of the protein (26). We used standard pulse sequences (27) to record steady-state [1H]-15N NOE

![FIGURE 3. Secondary chemical shifts (experimental minus random coil) for Hsp12.](image)

**TABLE 1**

Statistics of NMR structures of protein Hsp12 in SDS micelles and DPC at 40 °C

| Parameter                                             | SDS    | DPC (74–100) |
|-------------------------------------------------------|--------|--------------|
| Conformationally restricting distance constraints     |        |              |
| Intraresidue (i = j)                                   | 470    | 155          |
| Sequential (i = j)                                     | 376    | 104          |
| Medium range (i < (i – j) ≤ 5)                         | 282    | 63           |
| Long range ((i – j) > 5)                               | 0      | 0            |
| Total                                                 | 1,128  | 322          |
| Dihedral angle constraints                             |        |              |
| ϕ                                                     | 62     | 26           |
| ψ                                                     | 62     | 26           |
| Number of constraints per residue                      | 11.4   | 13.8         |
| Number of long range constraints per residue           | 0      | 0            |
| CYANA target function (Å)                              | 2.11 ± 0.16 | 1.52 ± 0.1   |
| Average r.m.s.d. to the mean CYANA coordinates (Å)     |        |              |
| Helix I, backbone heavy (F3–K18)                      | 0.08 ± 0.04 |              |
| Helix I, all heavy atoms (F3–K18)                     | 0.67 ± 0.15 |              |
| Helix II, backbone heavy (Y25–G32)                    | 0.17 ± 0.05 |              |
| Helix II, all heavy atoms (Y25–G32)                   | 0.66 ± 0.08 |              |
| Helix III, backbone heavy (Q94–S99)                   | 0.41 ± 0.14 |              |
| Helix III, all heavy atoms (Q94–S99)                  | 0.90 ± 0.19 |              |
| Helix IV, backbone heavy (D76–V100)                   | 0.22 ± 0.09 | 0.30 ± 0.14 |
| Helix IV, all heavy atoms (D76–V100)                  | 0.75 ± 0.11 | 0.83 ± 0.11 |
| PROCHECK raw score (ϕ and ψ/all dihedral angles) (34) | 0.21/–0.42 | 0.44/–0.24   |
| PROCHECK z-scores (ϕ and ψ/all dihedral angles)       | 1.14/–2.48 | 2.05/–1.42   |
| MOLPROBITY raw score/z-score (35)                     | 6.05/0.49 | 2.97/1.02    |

Ramachandran plot summary for ordered residues (%)

| Region                  | SDS    | DPC (74–100) |
|-------------------------|--------|--------------|
| Most favored regions    | 97.6   | 96.9         |
| Additionally allowed regions | 2.4 | 3.1          |
| Generously allowed regions | 0.0 | 0.0          |
| Disallowed regions      | 0.0    | 0.0          |

Average number of distance constraints violations per CYANA conformer (Å)

| Distance constraints violations | SDS    | DPC (74–100) |
|---------------------------------|--------|--------------|
| 0.2–0.5                         | 0.0    | 0.0          |
| >0.5                            | 0.0    | 0.0          |

Average number of dihedral angle constraint violations per CYANA conformer (º)

| Violations | SDS    | DPC (74–100) |
|------------|--------|--------------|
| 0          | 0      | 0            |

* Residues 9–19 (α-helix I), 25–44 (α-helix II), 52–61 (α-helix III), 74–100 (α-helix IV) considered by PSVS server.

* Residues 74–99 (α-helix IV) considered by PSVS server.
and 15N relaxation ($T_1$, $T_2$) data at 600 MHz. We acquired multiple interleaved NMR spectra with relaxation delays of 0, 50, 100, 150, 210, 280, 340, 440, 600, 800, and 1100 ms for $T_1$ and 10, 30, 50, 70, 90, 110, 130, 150, 170, 190, and 210 ms for $T_2$ determinations. The relaxation rates were calculated by least squares fitting of peak heights versus relaxation delay to a single exponential decay. The reported error estimates are standard deviations derived from fitting the data. Steady-state [1H]-15N NOE values were calculated from the ratio of peak heights in a pair of NMR spectra acquired with and without 3-s proton saturation. The signal-to-noise ratio in each spectrum was used to estimate the experimental uncertainty.

Structure Calculations and Analysis—We derived 1H-1H distance restraints from 15N-resolved three-dimensional 1H-1H NOESY and 13C-resolved three-dimensional 1H-1H NOESY spectra. We used TALOS+ software (28) to obtain backbone dihedral angle restraints ($\phi$ and $\psi$) from assigned 1H, 15N, 13C, 13C', and 13C chemical shifts and CYANA software version 3.0 (29) for automated NOEY peak assignments and structure calculations. We used MOLMOL (30) and PyMOL (31) software to calculate the root mean square deviation (r.m.s.d.) and for graphical analysis. We used the PSVS server (32) to check structure quality.

RESULTS AND DISCUSSION

Resonance Assignments and Secondary Structure of Hsp12—The 15N HSQC spectrum of Hsp12 (Fig. 1A) is characterized by peaks with uniform line width and signal intensity indicating that the protein was suitable for NMR study. However, the poor resonance dispersion in the 1H dimension implied that the protein is natively disordered, which was further confirmed by circular dichroism (Fig. 2), heteronuclear 1H-15N NOE experiments (see Fig. 5), and the lack of long range interactions in 1H-1H NOESY spectra. Traditional triple-resonance datasets were collected to make full sequence-specific backbone resonance assignment for the protein. Comparison of secondary 13C chemical shifts with random coil shifts further confirmed that the protein backbone is intrinsically disordered (Fig. 3). Accordingly, the TALOS+ program failed to predict any secondary structural elements. The assigned backbone and CB chemical shifts are available from BioMagResBank under accession code 17483.
Resonance Assignments and Secondary Structure of Hsp12_SDS and Hsp12_DPC—We collected a series of $^{15}\text{N}$ HSQC spectra upon adding increasing amounts of SDS or DPC micelles. In both cases, the spectra showed large chemical shift differences indicative of conformational change. However, Hsp12 in presence of SDS micelles at 40 °C (Fig. 1B) showed more uniform signal intensity and peak distribution than in the presence of DPC (Fig. 1C). The induction of helix by added SDS was evident from CD spectra (Fig. 2) and was confirmed by $^{13}\text{C}_\alpha$ secondary chemical shifts (Fig. 3B). Analysis of assigned chem-
Structural Characterization of Hsp12

Fig. 7. Effect of shift reagents on glycine $^1$H-$^15$N HSQC signals of Hsp12 in presence of 200 mM SDS. A, water-soluble shift reagent, 10 mM Gd(DTPA-BMA). B, no shift reagent (control). C, lipid-soluble shift reagent, 10 mM 16DSA. Spectra were recorded at 40 °C on a 900 MHz spectrometer equipped with a cryogenic probe.

Physical shifts by both TALOS+ and PINE indicated the presence of four $\alpha$-helices in presence of SDS and one $\alpha$-helix in presence of DPC. The assigned chemical shifts are available from BioMagResBank under accession codes 17482 (Hsp12_SDS) and 17948 (Hsp12_DPC).

Structure Determinations of Hsp12_SDS and Hsp12_DPC—We used advanced NMR techniques in attempting to solve the three-dimensional structure of Hsp12 in the presence of detergent micelles. In the presence of SDS, a total of 1126 distance constraints and 124 dihedral angle constraints derived from TALOS+ (Table 1) were used as input for CYANA structural calculations. The results confirmed the presence of four helical regions, but the software failed to generate a global fold because of the absence of long range NOEs indicating interactions between the $\alpha$-helices. The individual helices contained the characteristic NOE signatures of $\alpha$-helices, and by using these constraints we were able to derive a high quality model for the secondary structure. Fig. 4 shows the four well defined $\alpha$-helices consisting of residues Phe9-Lys18 (\alphaI), Tyr25-Gly42 (\alphaII), Gln54-Ser59 (\alphaIII), and Asp76-Val100 (\alphaIV). The r.m.s.d. for the backbone heavy atoms to the mean coordinates were 0.08 Å for \alphaI, 0.17 Å for \alphaII, 0.41 Å for \alphaIII, and 0.22 Å for \alphaIV (Table 1). Helix \alphaIII was not well defined as the others as indicated by its r.m.s.d. and $^{13}$C\* secondary shifts (Fig. 3). The coordinates of the structural models are available from the Protein Data Bank under accession code 2L9Q.

In the presence of DPC, Hsp12 gained a single helix that is similar to helix \alphaIV (Asp76-Val100). A total of 322 distance constraints and 52 dihedral angle constraints derived from TALOS+ were used to generate helix \alphaIV with the backbone r.m.s.d. of 0.31 Å (Fig. 4). The secondary shifts from the CA carbons additionally support the presence of a single helix (Fig. 3C). The coordinates of the structural models are available from the Protein Data Bank under the accession code 2LJI.

Backbone Dynamics of Hsp12, Hsp12_SDS, and Hsp12_DPC—Hsp12 in aqueous solution showed a uniform distribution of $T_1$, $T_2$ values, indicating that the protein backbone has uniform mobility throughout the protein sequence. The short $T_1$ values (∼400 ms), comparable $T_2$ values (∼300 ms), and negative $^{15}$N NOEs indicate that the protein backbone is highly dynamic (Fig. 5). Much of this mobility was lost upon the addition of SDS micelles as shown by the longer $T_1$ values (∼600 ms), shorter $T_2$ values (∼100 ms), and positive $^{15}$N NOEs (Fig. 5). The $T_1$, $T_2$, and $^{15}$N NOE values were uniform across the peptide sequence, apart from few N- and C-terminal residues and residues 65–75, which appear to form a loop. In HSP_DPC, positive $^{15}$N NOE values for the residues 74–100 (\alphaIV) additionally support the presence of single $\alpha$-helix (Fig. 5).

Paramagnetic Spin Relaxation Experiments—Spin relaxation agents provide tools for identifying membrane-associated and aqueous regions of proteins (33). The strategy behind the use of the three different paramagnetic spin labels is represented pictorially in Fig. 6. We compared $^1$H-$^15$N HSQC signal intensities in spectra of Hsp12_SDS without and with spin labels. Loss of signal intensity resulting from Gd(DTPA-BMA) identified residues exposed to the solvent. Loss of signal intensity resulting from 5DSA or 16DSA indicated residues facing the SDS micelles. The glycine region of $^1$H-$^15$N HSQC spectra (Fig. 7) clearly indicates the effect of paramagnetic spin relaxation agents. Amide resonances from residues Gly5, Gly55, Gly63, Gly71, Gly83, and Gly102, which were affected by Gd(DTPA-BMA), were unaffected by 16DSA. However, resonances from residues Gly10, Gly29, and Gly42, which were affected by 16DSA, were not affected by Gd(DTPA-BMA).

CONCLUSIONS

Our extensive structural characterization of Hsp12 by NMR, CD, and paramagnetic relaxation experiments show that Hsp12 is intrinsically disordered in aqueous solution but folds to generate four local $\alpha$-helices in the presence of SDS micelles or one local $\alpha$-helix in the presence of DPC. NMR-derived backbone relaxation parameters ($T_1$, $T_2$, and $^{15}$N NOE) showed a significant difference between the disordered and folded protein. The effects of added paramagnetic spin labels to Hsp12 in the presence of SDS micelles are consistent with the view that the protein associates with the plasma membrane. A recent, more qualitative, NMR investigation of Hsp12 in the presence of SDS micelles also found that the detergent induced formation of four $\alpha$-helices (20), but they did not extend their NMR study to other detergent micelles.
Acknowledgments—NMR data were collected at the National Magnetic Resonance Facility (Madison). We thank other staff members of the Center for Eukaryotic Structural Genomics who assisted in preparing the protein sample.

REFERENCES

1. Craig, E. A. (1993) Science 260, 1902–1903
2. Hendrick, J. P., and Hartl, F. U. (1993) Annu. Rev. Biochem. 62, 349–384
3. Martin, J., and Hartl, F. U. (1997) Curr. Opin. Struct. Biol. 7, 41–52
4. Lindquist, S., and Craig, E. A. (1988) Annu. Rev. Genet. 22, 631–677
5. Schlesinger, M. J. (1990) J. Biol. Chem. 265, 12111–12114
6. Tissières, A., Mitchell, H. K., and Tracy, U. M. (1974) J. Mol. Biol. 84, 389–398
7. Praekelt, U. M., and Meacock, P. A. (1990) Mol. Gen. Genet. 223, 97–106
8. Stone, R. L., Matarese, V., Magee, B. B., Magee, P. T., and Bernlohr, D. A. (1990) Gene 96, 171–176
9. Mtwisha, L., Brandt, W., McCready, S., and Lindsey, G. G. (1998) Plant Mol. Biol. 37, 513–521
10. Sales, K., Brandt, W., Rumbak, E., and Lindsey, G. (2000) Bba Biomembranes 1463, 267–278
11. Motshwene, P., Karreman, R., Kgari, G., Brandt, W., and Lindsey, G. (1998) Plant Mol. Biol. 37, 513–521
12. Varela, J. C., Praekelt, U. M., Meacock, P. A., Planta, R. J., and Mager, W. H. (2001) Mycobiology 155, 2021–2028
13. Schade, B., Jansen, G., Whiteway, M., Entian, K. D., and Thomas, D. Y. (2004) Mol. Biol. Cell 15, 5492–5502
14. Sano, F., Asakawa, N., Inoue, Y., and Sakurai, M. (1999) Cryobiology 39, 80–87
15. Park, J. I., Grant, C. M., Atfield, P. V., and Dawes, I. W. (1997) Appl. Environ. Microbiol. 63, 3818–3824
16. Sharma, S. C. (1997) FEMS Microbiol. Lett. 152, 11–15
17. Shanmuck, V. J., and Lindsey, G. G. (2008) Can. J. Microbiol. 54, 559–568
18. Welker, S., Rudolph, B., Frenzel, E., Hagn, F., Liebisb, G., Schmitz, G., Scheuring, J., Kerth, A., Blume, A., Weinkauf, S., Haslbeck, M., Kessler, H., and Buchner, J. (2010) Mol. Cell 39, 507–520
19. Tyler, R. C., Sreenath, H. K., Singh, S., Aceti, D. J., Bingman, C. A., Markley, J. L., and Fox, B. G. (2005) Protein Expr. Purif. 40, 268–278
20. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Prog. NMR Spectrosc. 34, 93–158
21. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
22. Bartels, C., Xia, T. H., Billeter, M., Güntert, P., and Wüthrich, K. (1995) J. Biomol. NMR 6, 1–10
23. Bahrami, A., Assadi, A. H., Markley, J. L., and Eghbalnia, H. R. (2009) PLoS Comput. Biol. 5, e1000307
24. Kay, L. E., Keifer, P., and Saarinen, T. (1992) J. Am. Chem. Soc. 114, 10663–10665
25. Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Kay, L. E. (1995) Biochemistry 34, 868–878
26. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) J. Biomol. NMR 44, 213–223
27. Güntert, P. (2004) Methods Mol. Biol. 278, 353–378
28. Laskowski, R. A., Rullmann, J. A., Kaptein, R., and Thornton, J. M. (1996) J. Mol. Graph. 14, 51–55
29. DeLano, W. L. (2010) The PyMOL Molecular Graphics System, version 1.3r1, Schrödinger, LLC, New York
30. Bhattacharya, A., Tejero, R., and Montelione, G. T. (2007) Proteins 66, 778–795
31. Hilty, C., Wider, G., Fernández, C., and Wüthrich, K. (2004) Chem. Bio. Chem. 5, 467–473
32. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arenedall, W. B., 3rd, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) Nucleic Acids Res. 35, W375–383