Structural insights into the dynamics and function of the C-terminus of the E. coli RNA chaperone Hfq

Journal Article

http://eprints.bbk.ac.uk/3320

Version: Published (Refereed)

Citation:


© 2011 Birkbeck College

Publisher version

All articles available through Birkbeck ePrints are protected by intellectual property law, including copyright law. Any use made of the contents should comply with the relevant law.
Structural insights into the dynamics and function of the E. coli RNA chaperone Hfq

Mads Beich-Frandsen1, Branislav Večerek2, Petr V. Konarev3, Björn Sjöblom1, Karin Kloiber1, Hermann Hämmerle2, Lukas Rajkowitsch2, Andrew J. Miles4, Georg Kontaxis1, B. A. Wallace4, Dimitri I. Svergun3, Robert Konrat1, Udo Bläsi2,* and Kristina Djinović-Carugo1,5,*

1Department of Structural and Computational Biology, Max F. Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, A-1030 Vienna, 2Department of Microbiology, Immunobiology and Genetics, Max F. Perutz Laboratories, University of Vienna, Dr. Bohrgasse 9, A-1030 Vienna, Austria, 3EMBL-Hamburg c/o DESY, Notkestrasse 85, D-22603, Hamburg, Germany, 4Department of Crystallography, Institute of Structural and Molecular Biology, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK and 5Department of Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Aškerčeva 5, 1000 Ljubljana, Slovenia

Received September 28, 2010; Revised December 18, 2010; Accepted December 22, 2010

ABSTRACT
The hexameric Escherichia coli RNA chaperone Hfq (HfqEc) is involved in riboregulation of target mRNAs by small trans-encoded RNAs. Hfq proteins of different bacteria comprise an evolutionarily conserved core, whereas the C-terminus is variable in length. Although the structure of the conserved core has been elucidated for several Hfq proteins, no structural information has yet been obtained for the C-terminus. Using bioinformatics, nuclear magnetic resonance spectroscopy, synchrotron radiation circular dichroism (SRCD) spectroscopy and small angle X-ray scattering we provide for the first time insights into the conformation and dynamic properties of the C-terminal extension of HfqEc. These studies indicate that the C-termini are flexible and extend laterally away from the hexameric core, displaying in this way features typical of intrinsically disordered proteins that facilitate intermolecular interactions. We identified a minimal, intrinsically disordered region of the C-terminus supporting the interactions with longer RNA fragments. This minimal region together with rest of the C-terminal extension provides a flexible moiety capable of tethering long and structurally diverse RNA molecules. Furthermore, SRCD spectroscopy supported the hypothesis that RNA fragments exceeding a certain length interact with the C-termini of HfqEc.

INTRODUCTION
In bacteria small trans-encoded regulatory RNAs (sRNAs) can modulate target gene expression in response to various stress conditions, a mechanism known as riboregulation. The majority of functionally characterized Escherichia coli sRNAs act as negative regulators by preventing ribosome loading onto the target mRNA through base-pairing with or in the vicinity of the ribosome binding site. As a result, the respective mRNA is prone to rapid decay [for review see refs (1–3)]. Positive regulation by sRNAs has been reported less frequently. In this case the sRNAs can act by an ‘anti-antisense’ mechanism and melt intramolecular secondary structures, which impede ribosome binding (4–6). Alternatively, an mRNA can act as a decoy for a sRNA leading to translation of another mRNA that is negatively controlled by this sRNA (7). Similarly, a sRNA can protect a homologous sRNA from degradation, which in turn results in activation of a target mRNA by the latter (8). A number of bacterial sRNAs associate with the global regulator Hfq (9) and often require this protein for function (10,11). Part of the Hfq requirement may be ascribed to Hfq-mediated stabilization of sRNAs (12–15).

*To whom correspondence should be addressed. Tel: +43 1 4277 52203; Fax: +43 1 4277 9522; Email: kristina.djinovic@univie.ac.at
Correspondence may also be addressed to Udo Bläsi. Tel: +43 1 4277 54609; Fax: +43 1 4277 9546; Email: udo.blaesi@univie.ac.at
Present address:
Lukas Rajkowitsch, Lexogen GmbH, Brunnerstrasse 69/Obj. 3, A-1230 Vienna, Austria.

© The Author(s) 2011. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
However, as many sRNAs display imperfect and non-contiguous target complementarity, the requirement for Hfq in riboregulation has mainly been attributed to its RNA chaperone function, which is believed to facilitate the interaction between the sRNA and the cognate mRNA (13,16–19).

Hfq belongs to the eukaryotic and archaeal family of Sm- and Sm-like proteins (13,19,20). Hfq orthologues have been found in a number of prokaryotes. The Hfq proteins of different organisms display an evolutionarily conserved core consisting of amino acid residues 7–66, whereas there is considerable variation at the C-terminal extension (20). The first X-ray structure of the Staphylococcus aureus Hfq protein (HfqSa) in complex with a short RNA oligonucleotide, was reported by Schumacher et al. (21). Since then 13 unique structures of Hfq-proteins from seven different organisms have been deposited in the Protein Data Bank [S. aureus (HfqSa; amino acids 1–77) 1KQ1, 1KQ2 (21); E. coli (HfqEc; amino acids 1–72) 1HK9 (22), 3GIB (amino acids 2–69) (23); Pseudomonas aeruginosa (HfqPa; amino acids 1–82) 1UIT, 1UIS (24), 3M4G, 3INZ (25); Methanococcus jannaschii (HfqMj; amino acids 1–71) 2QTX (26); Synechocystis sp. (HfqCys; amino acids 1–70) 3HFO; Anabaena sp. (HfqCAn; amino acids 1–72) 3HSB (28), and 3HSA (to be released)].

The study by Schumacher et al. (21) showed that HfqSa forms homo-hexamers with a central pore and that a short poly(U) oligoribonucleotide was bound in a circular conformation along the inner basic rim. A mutational analysis of poly(U) oligoribonucleotide was bound in a circular conformation and shape of the C-terminal extension. A mutational analysis by Mickulecky et al. (29) revealed binding of a poly(A) tract to the distal face of HfqEc (29). Unlike uridine-containing sequences, a recent structural study analyzed the dynamic properties, and to determine the conformation and shape of the C-terminal extension of HfqEc. Collectively, these studies revealed that the C-terminus is intrinsically disordered and flexible, which appears to contribute to the interactions of Hfq with RNA substrates.

### MATERIALS AND METHODS

#### Bacterial strains and growth conditions

The E. coli strains BL21 DE3 (Novagen), AM111 (MC4100 hfq::Ω) (34) and the corresponding F′ (lacP) variant (35) have been described. They were grown at 37°C in Luria–Bertani (LB) medium or in M9 medium supplemented with 0.2% glucose, 2 mM MgSO4, 0.1 mM CaCl2 and 10 μg/ml thiamine. For RpoS synthesis the cells were grown in LB medium at 22°C. Where appropriate, ampicillin (100 μg/ml), kanamycin (25 μg/ml), tetracycline (30 μg/ml) or chloramphenicol (15 μg/ml) was added to the medium.

#### Construction of plasmids and expression of hfq variants

The plasmids pUHS5 and pUH65 used for synthesis of full-length HfqEc and HfqEc65, respectively, have previously been described (18,36). In these plasmids, the corresponding hfq gene was placed under transcriptional control of the lac promoter in plasmid pUC19 (New England Biolabs). To avoid negative autoregulation (37), the 5′-UTR of the hfq gene was replaced by the ribosome binding site of phage T7 gene 10 using a PCR approach.

The plasmids pUH85 and pUH75 used for synthesis of HfqEc85 and HfqEc75 were constructed as follows. The forward primer (5′-GCCTCAGAATATAATAGTTTATCTTTAAGAAGGGAGATATACATGGCAATCTTTACAAGATCCGTTCCT-3′) and the reverse primer (5′-TTGGAATCTAATGTGTTACTGTGATGAG-3′) contained the sequence derived from plasmid pET22b (underlined; Novagen) including the XbaI site and the Shine and Dalgarno sequence of gene 10 abutted with the first 35 nucleotides of the hfq gene (given in bold). The reverse primers (5′-TTGGAATCTAATGTGTTACTGTGATGAG-3′) and (5′-TTGGAATCTAATGTGTTACTGTGATGAG-3′) bring two stop codons after the triplet encoding the Hfq-specific amino acids His85 and Ala75 as well as an EcoRI and BamHI site, respectively. The resulting PCR products were cleaved with XbaI and EcoRI/BamHI and then cloned into the same sites of plasmid pUC19 yielding plasmids pUH85 and pUH75, respectively.

The plasmids pProSA and pProBS used for purification of the S. aureus and B. subtilis Hfq proteins, respectively, have been described (18). For construction of plasmids pAH75 and pAH85, the corresponding hfq segments were sub-cloned as PvuII fragments from pUH75 and pUH85 and then inserted into the EcoRV and NruI sites of plasmid pACYC184. Construction of plasmids pAHfq and pAH65 is described (18).
Full-length HfqEc was produced in *E. coli* BL21 DE3 (Novagen) and the *S. aureus, B. subtilis* and truncated *E. coli* Hfq proteins were produced in the hfq deficient strain AM111F (lacI'). One liter of Luria–Bertani medium was supplemented with 100 µg/ml of ampicillin, 50 µg/ml kanamycin and 15 µg/ml tetracycline where appropriate. Protein for NMR-studies was obtained from *E. coli* Hfq proteins were produced in the (Novagen) and the variants thereof comprised four steps: (i) An initial Purification of Hfq proteins is described (38). The Hfq proteins are heat-stable, and the purification variants was induced by addition of 0.5 mM Isopropyl β-D-thiogalactopyranoside (IPTG) (Sigma). After 4 h the cells were harvested by centrifugation (4,000 g/15 min).

**In vitro synthesis of RNA**

For hfg126 mRNA synthesis plasmid pUfg126 (37) cleaved with AflIII was used as a template for in vitro transcription with the T7-MEGAshortscript kit (Ambion). The run-off transcripts were purified on 6% polyacrylamide-8M urea gels following standard procedures. The mRNA concentration was determined spectrophotometrically measuring absorption at 260 nm. The DsrA₃₄ RNA, comprising the nucleotides 26–59 of the DsrA-sequence was purchased from Dharmacon, USA.

**Purification of Hfq proteins**

The Hfq proteins are heat-stable, and the purification therefore involved an initial fractionation by heating and subsequent processing by FPLC (Akta, GE-Healthcare). The final purification-scheme for *E. coli* Hfq and variants thereof comprised four steps: (i) An initial washing step by Ni²⁺-affinity (HfqEc and HfqEc65) or hydrophobic interaction chromatography (HfqEc75, HfqEc65); (ii) a filtration over an anion-exchange column to remove nucleic acids; (iii) a concentrating step by respectively Ni²⁺-affinity or HIC; and (iv) a final step of size-exclusion chromatography. The detailed procedure is described in Supplementary Data. Purification of HfqBs and HfqSa was performed as previously described (18).

**Western blot analysis**

The RpoS, Hfq and L14 (loading control) protein levels were determined in strain AM111 harboring plasmids pACYC184 (control), pAHfq, pAH65, pAH75 and pAH85, respectively. The strains were grown at 22°C in LB medium to OD₆₀₀ of 0.5, at which time 2 ml aliquots were withdrawn, pelleted and boiled in protein sample buffer. Equal amounts of total protein were separated on 12% SDS–polyacrylamide gels and blotted to a nitrocellulose membrane. The blots were blocked with 5% dry milk in TBS buffer, and then probed with anti-RpoS (NeoClone, Madison), anti-Hfq (Pineda, Berlin) and anti-L 14 (provided by Dr I. Moll) antibodies to detect the RpoS, Hfq and L14 proteins, respectively. The antibody–antigen complex was visualized as described (38).

**FRET assays**

This method is described in more detail in refs (32,39). Two complementary, fluorophore-tagged RNA 21mers [Cy5-5’-AUGUGGAAAUCUCUAGCAGU-3’ (Cy5-21R⁺) and Cy3-5’-ACUGCUAGAUUUUCCA CAU-3’ (Cy3-21R⁻)] were used in the annealing experiment. The tagged RNA oligonucleotides were purchased from VBC-Biotech (Vienna, Austria). Using a first oligoribo nucleotidr. The reaction was performed in annealing buffer (50 mM Tris–HCl pH 7.5, 3 mM MgCl₂ and 1 mM DTT) at 37°C. The final concentration of the RNAs was 5 nM, in a sample volume of 40 µl. The reaction was allowed to proceed for 180 s, and with Cy3 excited, donor and acceptor dye fluorescence emissions were measured once every second. The time-resolved ratio of the fluorescence emissions (FRET index F_{Cy5}/F_{Cy3}) was normalized to 1 at t₀, and least-square fitted with Prism 4.03 (GraphPad Software Inc., San Diego, CA, USA) with the second-order reaction equation for equimolar initial reactant concentrations: $y = A [1 - (k_{obs} t + 1)^{-1}]$, where $y$ = fraction annealed, $k_{obs}$ = observed annealing reaction constant, $A$ = maximum reaction amplitude. The observed reaction constant $k_{obs}$ shown in Figure 1A was calculated as the average of three individually fitted reactions.

**SRCD Spectroscopy**

SRCD spectra were collected on the CD1 beamline at the ISA facility, University of Aarhus, Denmark. For each protein, the extinction coefficient at 280 nm [HfqEc ~0.34, HfqEc₆₅ ~0.41, HfqEc₇₅ ~0.30, HfqEc₆₅ ~0.35 [(mg/ml⁻¹ × cm⁻¹)] was calculated from its amino acid sequence (http://www.scripps.edu/~cdp/nam/protcalc.html) and the concentration derived from the A₂₈₀ measured on a Nanodrop 1000 spectrophotometer immediately prior to the SRCD spectrum being acquired. For the complexes with RNA, the protein concentration was assessed prior to RNA addition to allow for a 1:1 molar ratio of Hfq hexamer to RNA. The protein and RNA was mixed at elevated temperature (~60°C), and incubated for 5 h at 65°C, and then allowed to slowly cool to room temperature. Subsequently the complexes were centrifuged for 1 h at 10,000 g, at room temperature.

Each protein sample (concentrations of ~5 mg/ml) was loaded into a 0.015 mm path length quartz Suprasil demountable cell (Hellma Ltd) and three spectra were collected at 20°C over the wavelength range from 280 to 170 nm, with a 1 nm step size and a 2 s dwell time. Data processing was carried out using the CDTool software (40). Replicas scans were averaged, subtracted from the average of three baseline spectra (either buffer or buffer plus RNA), smoothed with a Savitsky–Golay filter, and calibrated against a spectrum of camphoursulphonic acid that was obtained at the beginning of the data collection (41). The spectra were converted to delta epsilon units using mean residue weight.
values [calculated as molecular mass/number of residue\(^{-1}\)] of 110.6, 114.3, 113.7 and 112.5 for the H\(\text{f}_{1}\)Ec, H\(\text{f}_{2}\)Ec85, H\(\text{f}_{3}\)Ec75, and H\(\text{f}_{4}\)EcE85 constructs, respectively, based on molecular weights of 11166.3, 7317.5, 8417.7, 9415.7, and numbers of residues of 102, 64, 75, 85. The low wavelength cutoff of the data was determined as previously described (42). Data were analysed with the DichroWeb analysis server (43) and are reported as the averaged results (+1 SD) from the CONTINLL (44,45), SELCON3 (46) and CDDSTR (47) algorithms, using the SP175 reference dataset (48). The NRMSD is a goodness-of-fit parameter (49) calculated for the CONTINLL-calculated structure and the experimental data.

NMR

All NMR experiments were performed at 310 K on a Varian Inova 600 MHz spectrometer equipped with 5 mm triple resonance cryo-probe and pulsed field gradients, and a Varian Inova 800 MHz spectrometer equipped with 5 mm conventional triple resonance probe equipped with pulsed field gradients. NMR spectra were processed with NMRePipe (50) and analyzed with Sparky software. The protein samples for NMR-experiments were all prepared by size-exclusion chromatography in 50 mM Na-PO\(_4\) pH ~7.2, 200 mM NaCl and concentrated to ~1 mM H\(\text{f}_{4}\)EcE85 (with respect to the monomer). Samples were supplemented with ~10\% (v/v) \(\text{H}_{2}\)O to provide the deuterium signal for the field-frequency lock, as well as 0.1–0.2\% (w/v) NaN\(_3\) to inhibit bacterial growth. Backbone signal assignment for the C-terminally truncated mutant H\(\text{f}_{4}\)EcE85 was obtained by a suite of standard (sensitivity enhanced) three-dimensional (3D) triple resonance experiments, acquired pairwise: HNCA (51)/HN(CO)CA (52), HNCACB (53), HNCO (54) were recorded for sequential backbone chemical shift assignment, \(^{15}\)N relaxation times \((T_1, T_2)\) were measured using well-known gradient sensitivity-enhanced 2D methods with \(^1\)H detection (55,56).

SAXS

Small angle X-ray scattering experiments were acquired at beamline X33 (57,58) at the DORIS III synchrotron storage-ring (DESY, Hamburg, Germany), using a MAR345 image-plate detector and X-ray wavelength of \(\lambda = 1.5\,\text{Å}\), with a sample to detector distance of 2.7 m. This setup covers a range of momentum transfer \(0.12 < s < 0.45\,\text{nm}^{-1}\) \((s = 4\pi\text{sin}θ/\lambda, \text{where } 2\theta \text{ is the scattering angle})\). The data were processed in the ATSAS program-package (59). Data were acquired at 37\(^\circ\)C from four concentrations of the full-length H\(\text{f}_{4}\)Ec (18.5, 9.3, 4.6, 2.3 mg/ml), and similar concentration of the C-terminally truncated core-construct H\(\text{f}_{4}\)EcE85 (19.0, 9.5, 4.8, 2.4 mg/ml). Protein concentration was measured with 2 \(μl\) samples on a Nanodrop 1000 UV/vis spectrophotometer at 280 nm, immediately prior to data-acquisition. The data were averaged and normalized to the intensity of the incident beam prior to subtraction of buffer scattering. The difference data curves were scaled, merged and initial analysis performed using PRIMUS (60). Invariants were derived by standard approaches for comparison (61). The intrinsic scattering \((I_0)\) and radius of gyration \((R_g)\) was derived by Guinier analysis for the data range \(s < 1.3/R_g\), approximating \(I(s) = I_0 \exp(-s R_g)^2/3\). The reverse transform program GNOM (62) was employed for calculation of the distance-distribution function, under the assumption of correct estimation of the particle maximum dimension \((D_{\text{max}})\). The program also outputs a value for \(I_0\) and \(R_g\). The particle excluded volume was calculated from the Porod Equation (63):

\[
V = 2\pi^2 I_0/\int_0^\infty s^2 I_{\text{exp}}(s)\,ds,
\]

where \(I_{\text{exp}}(s)\) is the experimental data.

In addition, the program Dammif (64) run without symmetry restraints outputs the particle excluded volume. The solute molecular mass was further evaluated by the standard approach of comparison of the intrinsic forward scattering \((I_0)\) with that recorded from a reference solution of known molecular mass and concentration. In the present study a single sample of 3.68 mg/ml bovine serum albumin \((\text{Mr} = 66.4\,\text{kDa})\) was used.

Low-resolution shape-models of the SAXS data were generated \(ab\ initio\) using the program Dammif (64). The program represents the protein as a collection of densely packed dummy-residues inside a sphere with the diameter \(D_{\text{MAX}}\). Each dummy-residue belongs either to the particle or the solvent, and the shape is described by a binary string of length \(M\). Starting from a random string, simulated annealing is employed to search for a compact model that fits the shape scattering curve \(I(s)\) to minimize discrepancy \(\chi^2\):

\[
\chi^2 = \frac{1}{N-1} \sum_j \left[\frac{I(s_j) - cI_{\text{calc}}(s_j)}{\sigma(s_j)}\right]^2,
\]

where \(N\) is the number of experimental points, \(c\) is a scaling factor and \(I_{\text{calc}}(s_j)\) and \(\sigma(s_j)\) are the calculated intensity and the experimental error at the momentum transfer \(s_j\), respectively.

The program BUNCH (65) employs a combined rigid body and \(ab\ initio\) modeling approach to search for a spatial arrangement of domains with known high-resolution structure, and probable conformations of flexible domains represented as dummy-residues. Utilizing the known structure of a H\(\text{f}_{4}\)EcE85 monomer (amino acids 6–65), the program was run employing a 6-fold symmetry retaining the hexameric organization of the H\(\text{f}\)q core. Theoretical scattering from hexameric H\(\text{f}\)q core encompassing residues 6–65 (PDBid: 1HK9) was calculated using the program CRY SOL (66). The C-termini (6 × 37 residues) were represented as chains of dummy residues connected to Ser65 of H\(\text{f}\)q. The program employs a simulated annealing procedure to manipulate the local conformation of the chains representing the termini, to minimize the \(\chi^2\)-discrepancy as described in
Equation (2). The program allows for specifying restraints for the conformation of the flexible chains. In present study, the disordered nature of the C-terminus was represented by lowering of the restraint on the dihedral angle between adjacent dummy residues.

In order to test the packing of the C-terminus into the groove formed by amino acids 35–55 of one subunit and the β-sheet of the adjacent subunit as observed in the crystal structure of HfqEc65 (amino acids 6–65), we calculated rigid body models using the program BUNCH applying distance restraints on His70. When placing His70 in the groove, to force the C-termini into the groove we observe a decrease in the quality of the fit. When placing His70 on the outside of the N-terminal helix, to constrain the C-terminus in the vicinity of the conserved patch of charged residues present in Hfq (22), no significant change to the quality of the fit was observed. All graphical representations were made using MacPymol (67).

RESULTS

Annealing deficiency of Hfq proteins naturally devoid of a C-terminus and of C-terminally truncated E. coli Hfq variants

We have recently shown using fluorescence resonance energy transfer assays (FRET) that full-length HfqEc stimulated annealing of two complementary RNA-oligonucleotides with a rate constant that was ~3-fold higher than that observed for HfqEc65, comprising only the evolutionarily conserved core region of Hfq (amino acids 1–65) (18). When two non-complementary RNA oligonucleotides were used, no significant increase in the FRET signal was discernable in the presence of HfqEc65, whereas an increase in FRET was observed for HfqEc. In contrast to HfqEc, the dual rate binding constant \( k_{obs} \) for HfqEc65 was ~10-fold lower, suggesting that HfqEc65 is severely impaired in annealing of two RNA oligonucleotides (18).

To test whether Hfq proteins naturally lacking a C-terminal extension are deficient or impaired in annealing activity, FRET assays were performed with the Hfq proteins of B. subtilis (HfqBs; 73 amino acids) and S. aureus (HfqSa; 77 amino acids). When compared to the absence of protein and similarly to HfqEc65, neither HfqBs nor HfqSa stimulated annealing of the complementary oligonucleotides (Figure 1A), again suggesting that the evolutionarily conserved core is not sufficient to support duplex formation of two RNAs.

With the aim to identify sub-sequences in the C-terminus of HfqEc that support RNA binding, we tested the ability of two C-terminally truncated HfqEc variants, HfqEc75 (amino acids 1–75) and HfqEc85 (amino acids 1–85) for their capacity to mediate low temperature annealing deficiency of Hfq proteins naturally devoid of a C-terminus and of C-terminally truncated E. coli Hfq variants. 5'-end were annealed at 37°C in the absence or presence of 100 nM protein. Relative fluorescence resonance energy transfer was calculated as the ratio of acceptor to donor fluorescence (FCy5/FCy3) as described in ‘Materials and Methods’ section. The time-resolved curves were least-square fitted with the second-order reaction equation for equimolar initial reactant concentrations: \( y = \frac{A}{1 + \left( k_{obs} \ t + 1 \right)^{-1}} \), where \( y \) = fraction annealed, \( k_{obs} \) = observed annealing reaction constant, \( A \) = maximum reaction amplitude. The reaction rate \( k_{obs} \) was calculated from the average of three independent experiments. (B) RpO8 protein synthesis in the E. coli hfq- strain AM111 harboring plasmids pACYC184 (control, lane 1), pAH65 (encoding HfqEc65, lane 2), pAH75 (encoding HfqEc75, lane 3), pAH85 (encoding HfqEc85, lane 4) and pAHfq (encoding full-length HfqEc, lane 5), respectively. The western-blot analysis was carried out with anti-RpO8, anti-Hfq antibodies and with antibodies against ribosomal protein L14 (loading control) using equal amounts of total cellular protein as described in ‘Materials and Methods section’. Only the relevant sections of the western-blot showing the RpO8-, Hfq- and L14-specific signals are shown.
Bioinformatics analyses predict an intrinsically disordered, solvent exposed C-terminus of HfqEc

The C-terminus of HfqEc (amino acids 66–102) contains a low portion (17%) of ‘order promoting’ residues (Asn, Cys, Ile, Leu, Phe, Trp, Tyr and Val), and a high fraction (64%) of ‘disorder promoting’ residues (Ala, Arg, Gln, Glu, Gly, Lys, Pro and Ser). As observed in several bioinformatic studies (69,70) this is a fingerprint of intrinsically disordered proteins. For comparison, in folded globular proteins, these values are 36% for ‘order’ and 47% for ‘disorder promoting’ residues, which gives a ratio of ~0.76. Interestingly, in the first half of the HfqEc C-terminus (amino acids 68–85) this ratio is ~0.63, whereas for the last half we calculate ~0.08. Overall this indicated that the C-terminal extension, which was indispensable for function in the assays used in this study (Figure 1), is predicted to be disordered.

The metaPrDOS algorithm, which generates disorder predictions based on a consensus principle between the individual predictors PrDOS, DISOPRED2, DisEMBL, DISPROT-VSL2P, DISpro and IUPred (http://prdos.hgc.jp/meta) (71), predicts the C-terminal moiety comprising residues 69–102, to be intrinsically disordered, with high signals (above 0.70) for amino acids 73–102 (Supplementary Figure S1). Secondary-structure prediction servers Jpred (72) and PSIPRED (73) do not detect secondary-structure elements in the C-terminal segment (66–102) of HfqEc (data not shown).

In addition, a protein meta-structural analysis was performed, which extracts two parameters for each residue on the basis of the protein primary sequence, one related to compactness and the other to secondary structure (74). The compactness parameter ($C_i$) provides a quantitative measure of the residue embedded in the protein tertiary structure. High values (>300) are found for residues in the buried core of globular domains, whereas low values correspond to surface exposed or conformationally flexible residues. The overall $C_i$ values for the C-terminal segment (amino acids 70–102) are lower than for the core of Hfq, suggesting that this region is conformationally flexible and surface exposed. In addition, the compactness of the Hfq-core region and that of the C-terminus were calculated separately. The average compactness value of the Hfq core region increased (from 292 of the full-length protein) to about 337, whereas the corresponding value for the C-terminal extension alone decreased to 170, indicating that the C-terminus is less compactly folded when compared to the Hfq core. The conclusion that the C-terminus is largely disordered is further supported by the fact that a similar average compactness value ($C_{av}$) were found for a set of experimentally characterized intrinsically unstructured proteins [Myc C-terminal domain (75): $C_{av} = 131$; osteopontin (76): $C_{av} = 181$; α-synuclein (77): $C_{av} = 255$; TAU (78): $C_{av} = 224$; Sic1 (79): $C_{av} = 229$].

This bioinformatic analysis (Figure 2; Supplementary Figure S1) together with the 3D structure of Hfq (22), in which residues 65–71 are found in an extended conformation leaning on the body Hfq core, suggest that the segment encompassing amino acid residues 65 up to 75 displays propensity to be structured, while this tendency is much less pronounced for the rest of the C-terminal extension with high indications for being intrinsically disordered. These features of HfqEc resemble intrinsically disordered protein sequences, which show a

**Figure 2.** Meta-structure analysis of HfqEc. Compactness values ($C_i$) are plotted with respect to residue position. A high value indicates a buried residue (the average $C_i$ of all residues in the Protein Data Bank is 300). Inserted in upper-left is the known secondary-structure content of Hfq (amino acids 6–65; PDBid: 1HK9) generated by ESPript (http://esprit.ibcp.fr/ESPRip/ESPRip/). TT denotes β-turn and η1 the small 310-helix. Bottom, primary amino acid sequence of HfqEc.
propensity to form local elements and/or contain pre-
formed elements of secondary structure or hydrophobic clusters (80,81).

Nuclear magnetic resonance and synchrotron radiation circular dichroism spectroscopy reveal the unstructured nature of the C-terminal extension of HfqEc

To obtain more information on the structural and dynamic properties of the C-terminus of HfqEc in solution we employed nuclear magnetic resonance spectroscopy (NMR). A series of NMR-spectra were recorded for full-length HfqEc and for HfqEc65. All NMR experiments for sequential signal assignments were conducted with HfqEc65, which—due to its lower molecular weight—was more amenable for NMR studies. Figure 3A shows the fingerprint region of the 15N-HSQC spectra of HfqEc65 superimposed on the 15N-HSQC spectra of full-length HfqEc. The spectrum of HfqEc65 is well resolved and sequential backbone signal assignment was possible for residues 6–65. Residues 1–5 were not observed possibly due to exchange broadening, suggesting some conformational flexibility of these residues. The spectra of HfqEc65 and full-length HfqEc overlapped largely, which allowed for assignment of the HfqEc spectra by simple peak comparison. However, many of the observed peaks corresponding to well-structured residues in the core of the protein show severe line broadening in the HfqEc spectrum due to fast relaxation. A plot of the 15N–1H chemical shift differences between HfqEc65 and full-length HfqEc is shown in Figure 3B. Significant chemical shift differences were observed for residues in the region preceding the C-terminal segment (amino acids 45–55) and at the end of the N-terminal α-helix (residues Phe11, Leu15, Arg17, Val20), indicating that the presence of the C-terminus affects residues in these regions. In the 3D structure from X-ray studies performed by Sauter et al. (22), the segment encompassing amino acids 65–70 was found to pack in an extended conformation against the core of the Hfq hexamer, in a groove formed by the N-terminal α-helix of one domain and β-sheets formed by amino acids 34–55 of the adjacent subunit. A conserved patch of charged residues comprised by Arg16, Arg17, Glu18 and Arg19 is present in Hfq. Interestingly, amino acids substitutions of Arg17 in HfqEc compromised phage Qβ replication (82), suggesting that this region is of functional importance.

The spectrum of full-length HfqEc shows additional intense peaks, mainly in the region of 7.5–8.5 ppm (in the 1H dimension), corresponding to the partially disordered C-terminus. Interestingly, in the region of the 15N-HSQC-spectra expected for Gly residues in a random-coil conformation, we did not observe any of the four extra glycine residues present in the C-terminus, indicating that they are indeed not fully unstructured. One possible explanation for their disappearance is that they are subjected to an additional dynamic process, leading to further line broadening, which could result from conformational exchange between a partially structured and a fully unstructured state. Another cause for disappearance of correlation peaks is rapid intermolecular exchange of amide protons with bulk water. This exchange process is particularly facilitated in disordered and conformationally flexible parts of the protein lacking stabilizing backbone hydrogen bonds.

In order to study the secondary structure composition of the C-terminal extension of HfqEc in solution, we used synchrotron radiation circular dichroism (SRCD) spectroscopy. The additional data in the low-wavelength region accessible by this technique relative to that attainable in a conventional CD instrument is important for detecting natively-disordered types of secondary structure. The spectra of the protein constructs HfqEc65 (amino acids 1–65), HfqEc75 (amino acids 1–75), HfqEc85 (amino acids 1–85) and HfqEc (full-length, amino acids 1–102) collected under the same experimental conditions are shown in Figure 4.

The spectrum of the HfqEc65 is typical for a mixed α/β protein and the calculated secondary structure based on these data produced values close to those calculated from the crystal structure (Table 1). When the HfqEc65 protein is extended by 10 residues (HfqEc75) there is a slight increase in spectral magnitude at both 222 and 190 nm, generally indicative of an increase in helix content, which is reflected in the analysis (Table 1) as a slight increase (2%) in the calculated helical content. This may be explained by stabilization of the N-terminal α-helix by the segment 65–75 through a similar packing arrangement observed in the crystal structure of E. coli Hfq (amino acid residues 1–72 PDBid: 1HK9) (22), giving rise to increased ellipticity. This is in good agreement with the NMR observations showing that this C-terminal stretch of residues interacts with the N-terminal α-helix.

In general, with each additional extension to the C-terminus (i.e. HfqEc85 and full-length HfqEc) there is a reduction in magnitude of the spectrum combined with a blue shift of the peak located at ~208 nm. This is indicative of an increase in the contribution of natively disordered-type structures to the spectrum, consistent with the assumption in the ‘theoretical’ values that the additional C-terminal residues are almost entirely unstructured; the percentage of ‘other’ increases as the number of residues are added.

The overall decrease in helix percentage as more residues are added—up to the full-length construct is caused by the original and constant number of helical residues representing a lower proportion of the total molecule in these constructs, not because the helical residues become unstructured. On the other hand, the percentage of ‘experimental’ β-structure in all examined variants is relatively constant at ~33%, while the expected ‘theoretical’ values decrease, as this is calculated assuming the additional regions in the constructs to be disordered. This suggests that the additional C-terminal residues contribute to some extent to an increase in the number of β-structure residues in HfqEc, which is in agreement with results from FTIR spectroscopy studies (83) showing the C-terminus to increase the overall β-structure content of the protein.
Figure 3. Comparative NMR analysis of HfqEc<sub>65</sub> and full-length HfqEc. (A) Overlay of $^{15}\text{N}$-HSQC spectra of HfqEc<sub>65</sub> [black (positive) and blue (negative) contours] and full-length HfqEc [red (positive) and green (negative) contours] demonstrating the unstructured nature of the C-terminal extension, with multiple intense peaks occurring in the 7.5–8.5 ppm region of the $^1\text{H}$-dimension. The glycine region in the upper part of the HfqEc spectra (yellow box) is lacking the expected four peaks for the four glycine residues of the C-terminus. (B) Chemical $^1\text{H}$-$^{15}\text{N}$ shift differences calculated for assigned peaks plotted against residue positions. Chemical shift differences >0.5 ppm are considered significant. The insert shows a surface representation of the proximal face and side of E. coli Hfq (amino acids 6-65) hexamer. Chemical $^1\text{H}$-$^{15}\text{N}$ shift differences between HfqEc<sub>65</sub> and full-length HfqEc are color-coded from blue (zero) to red (largest shift). Residues which could not be unambiguously assigned in the full-length HfqEc spectra are in blue. ‘C’ indicates the C-terminal Ser-65 of HfqEc<sub>65</sub>, noted in pink.
Solution structure, hydrodynamic properties and shape of the C-terminus of full-length HfqEc

To further study the behavior of Hfq in solution, we employed a series of solution small angle X-ray scattering (SAXS) experiments in combination with NMR spectroscopy measurements of the transverse relaxation rate of Hfq. Figure 5 shows the fit between experimental X-ray scattering data and the calculated scattering curve (66) for residues 6–65 of Hfq (PDBid: 1HK9) (22), confirming the close agreement between crystal and solution structure of the hexameric Hfq core. By modeling the five missing N-terminal residues, we observe only a marginal improvement ($\chi^2 = 1.55$), confirming a modest contribution to the overall scattering of these 6–5 = 30 unstructured residues in HfqEc65.

For HfqEc65 we estimated a molecular mass of 44 ± 5 kDa and for full-length HfqEc 58 ± 7 kDa by normalization of the intrinsic scattering ($I_0$) to a standard sample of bovine serum albumin (BSA) (84), which agrees with the theoretical values for hexamers of 43.2 and 67.2 kDa, respectively. We estimate the solute excluded volume (Pord’s volume) for HfqEc65 $(70 \pm 5) \times 10^3$ Å$^3$ and for full-length HfqEc $(110 \pm 8) \times 10^3$ Å$^3$, further indicating that both proteins adopt their native hexameric quaternary structure. In addition, Guinier analysis established for HfqEc65 an $R_g$ of $(24.2 \pm 0.5)$ Å, in good agreement with $R_g$ of 23.5 Å calculated for the crystal structure of HfqEc65. Comparatively, for full-length HfqEc $R_g$ was $(32.3 \pm 0.5)$ Å (60). The estimated largest dimension ($D_{\text{MAX}}$) of the particle was for HfqEc65 $(68 \pm 5)$ Å and for HfqEc $(112 \pm 8)$ Å (62).

By NMR $^{15}$N-relaxation measurements we measured for HfqEc65 (43.2 kDa) a $T_2 = 40.0 \pm 2.1$ ms ($R_2 = 25.0 \pm 1.3 s^{-1}$). These homogeneous $^{15}$N-relaxation data indicate that the core domain of Hfq is uniformly rigid and does not undergo large amplitude fluctuations on timescales relevant for NMR.

For full-length HfqEc (67.2 kDa), we measured an extremely fast $^{15}$N relaxation of $T_2 = 5–10$ ms for the signals of the core-region. Compared to HfqEc65, HfqEc exhibits

![Figure 4. SRCD spectra of Hfq constructs. HfqEc65 (thin dotted line), HfqEc75 (thick dotted line), HfqEc85 (thin solid line), full-length HfqEc (thick solid line). The error bars represent one standard deviation between replicate scans.](image)

Table 1. Secondary structure content of E. coli Hfq constructs and their complexes with RNA, evaluated from SRCD spectra using the DichroWeb analysis server (43)

<table>
<thead>
<tr>
<th>Protein</th>
<th>RNA</th>
<th>Theoretical %</th>
<th>Experimental %</th>
<th>NRMSD$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\alpha$-Helix</td>
<td>$\beta$-Strand</td>
<td>Other</td>
</tr>
<tr>
<td>HfqEc65</td>
<td></td>
<td>11</td>
<td>21</td>
<td>68</td>
</tr>
<tr>
<td>HfqEc85</td>
<td></td>
<td>13</td>
<td>25</td>
<td>62</td>
</tr>
<tr>
<td>HfqEc75</td>
<td></td>
<td>15</td>
<td>28</td>
<td>57</td>
</tr>
<tr>
<td>HfqEc65</td>
<td></td>
<td>17</td>
<td>32</td>
<td>51</td>
</tr>
<tr>
<td>HfqEc65</td>
<td>hfg126</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HfqEc65</td>
<td>sodB53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HfqEc</td>
<td>DsrA34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HfqEc65</td>
<td>DsrA34</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Bold values listed under theoretical correspond to the crystal structure of HfqEc65; other ‘theoretical’ values are calculated based on the crystal structure and assuming the extension residues are disordered.

$^b$Normalized root mean-square deviation between calculated and experimental spectra using the CONTINLL algorithm.
considerably reduced $^{15}$N-transverse relaxation times ($T_2 = 5–10$ ms). Consequently the precision of the $T_2$ measurement is much lower due to the resulting lower $S/N$-ratio. Interestingly, there is a larger residue by residue variation of $^{15}$N $T_2$ relaxation times in HfqEc than in HfqEc65, which is an indication of a more anisotropic behavior due to the extended C-termini. The observed relaxation rate is much faster than what could be expected for the increase in molecular mass, indicating an unusual increase in the hydrodynamic radius. For a perfect solid sphere, the radius of gyration ($R_g$) is only slightly smaller than the hydrodynamic radius ($R_h$), whereas increased discrepancy between the two reflects an anisometric structure of the particle. With a relation of 1.33 between the $R_h$ of HfqEc65 and HfqEc from SAXS measurements, reflecting the expected increase in molecular weight, the relation of 4–8 between the relaxation rates measured by NMR (which are directly proportional to $R_h$) indicates a pronounced anisometric structure of HfqEc. A possible explanation for these observations is that the partially unstructured C-termini are laterally extended into the surrounding solution. Thus the overall molecular shape is ‘puffed up’, causing the molecule to appear much larger than a comparable well-folded and tightly packed globular protein.

To validate this hypothesis, the observed $^{15}$N $T_2$ relaxation properties were modeled through simulation of the hydrodynamic properties of HfqEc65 and HfqEc using their molecular coordinates as input. For HfqEc65 the coordinates from the X-ray crystal structure of Hfq of *E. coli* (amino acid residues 6–65, PDBid: 1HK9) (22) were used, onto which the N-termini had been added as unstructured. For full-length HfqEc a representative all-atom ensemble was created by restrained simulated annealing using the X-plor/CNS software system (85) constraining the backbone dihedral angles of the C-terminus to the $\beta$-region of the Ramachandran plot and applying the observed radius of gyration as overall restraint.

Using the HYDRONMR software (86) for simulation of relaxation times both protein systems were very well reproduced in the simulation and good agreement was obtained between experimental and simulated relaxation

---

**Figure 5.** Overlay of experimental X-ray scattering data from HfqEc65 (black dots), on the calculated scattering curves for Hfq crystal structure (PDB: 1HK9, aa 6–65, red curve) and the complete model of HfqEc65 (aa 1-65, green curve) where N-terminal residues were restored using the Bunch program. Inserted is the pair-distribution function for HfqEc65 (grey).
behavior. For HfqEc\textsubscript{65} \textsuperscript{15}N transverse relaxation times of $T_2 = 35.0 \pm 2.0\text{ ms}$ ($R_2 = 28.6 \pm 1.6\text{ s}^{-1}$) were predicted; for the ensemble of full-length Hfq\textsubscript{Ec} models an average value of $^{15}$N $T_2 = 9.7 \pm 2.0\text{ ms}$ ($R_2 = 103.1 \pm 21.3\text{ s}^{-1}$) was found in the simulation.

Implying 6-fold symmetry the low-resolution shape of Hfq\textsubscript{Ec} was constructed \textit{ab initio} in the program Dammif (64). As shown in Figure 6A, the C-termini extend away from the core of the protein in a propeller like shape. The dimensions of the hexameric core of the \textit{ab initio} model (diameter $\sim$72 Å, thickness $\sim$35 Å) are larger than those derived from the crystal structure (diameter $\sim$65 Å, thickness $\sim$28 Å) (22), which can be ascribed to the presence of the termini. The output $R_g$ 32.6 Å by Dammif is in agreement with the $R_g$ derived from the Guinier-analysis (32.3 Å), and the partial specific volume of the protein

**Figure 6.** Shape models of Hfq\textsubscript{Ec}. (A) Surface representation of \textit{ab initio} model. (B) Calculated scattering curve of \textit{ab initio} model (blue) overlaid with experimental X-ray scattering data, inserted is the pair-distribution function for Hfq\textsubscript{Ec65} (grey) and full-length Hfq\textsubscript{Ec} (blue). (C) Superimposition of the three best rigid-body Hfq\textsubscript{Ec} models from three iterations of modeling (red, gold and green). In the center, the crystal structure of the Hfq\textsubscript{Ec} (amino acids 6–65) hexameric core is represented as grey ribbons and the N-termini as grey spheres. The \textit{ab initio} model is represented as blue spheres in the background. (D) Calculated scattering curves to a maximum range of momentum transfer ($s$) of 0.4 Å\textsuperscript{-1} for the three best rigid-body models (corresponding in red, gold and green).
were indistinguishable (Figure 7), the known binding conserved hexameric core between the modeling. Clear agreement was observed for the shape of the C-termini. By adjusting restraints in the observed for the C-termini. Iterative runs of Bunch (29), apparently does not affect the structure of the determined by mental data for the three best rigid-body models, as the Guinier plot. Figure 6D shows the fit with the experimental data on either the proximal or distal face of the hexamer. The models with the C-termini completely extended towards the program Bunch (65). We first built the N-terminal residues onto the crystal structure of HfqEc65 (Figure 5), and proceeded to utilize the outcome model as a rigid-body in reconstruction of the C-terminal domains of full-length HfqEc, again implying 6-fold symmetry in the particle. Figure 6C shows the superposition of the model of HfqEc from three iterations of rigid-body modeling. Clear agreement was observed for the conserved hexameric core between the ab initio and rigid-body models, whereas some variability was observed for the C-termini. Iterative runs of Bunch (×1000) generated models with low variability in the shape of the C-termini. By adjusting restraints in the Bunch-program we could fit the experimental data to \( \chi^2 \sim 1.65-1.75 \). We observed bad fits (\( \chi^2 > 2.3 \)) for models with the C-termini completely extended towards either the proximal or distal face of the hexamer. The calculated \( R_g \) 31.8 Å (average of three best fits) for the models is in good agreement though slightly smaller than the corresponding value of ~32.3 Å derived from the Guinier plot. Figure 6D shows the fit with the experimental data for the three best rigid-body models, as determined by \( \chi^2 \)-values generated by the program Bunch. These models fit the experimental data to a range of momentum transfer of 0.35 Å\(^{-1}\), and indicate that the C-termini extend laterally away from the hexameric core. The overall trend appears to be a wave-like shape, with a characteristic loop in the beginning of the C-termini and some variability in the tip. Interestingly, this wave-like behavior is also seen in the meta-structure analysis, where a \( \beta \)-turn like structural element is found in the disordered C-terminus (Figure 2).

**DISCUSSION**

In this integrative study on the dynamics and structure of the C-terminus of the Hfq protein from *E. coli*, we used bioinformatics and biophysical methods combined with molecular and structural biology. In brief, the different methods collectively revealed that the C-termini of HfqEc are intrinsically disordered and that they extend laterally away from the doughnut shaped core of the hexamer. Such flexible, intrinsically disordered regions are believed to provide conformational fluctuations, which can facilitate intermolecular interactions (33). Intrinsically disordered protein sequences are known to contain preformed, structured recognition motifs with functional implications in a hierarchical model of binding (88). Conversely, the unstructured nature of a binding motif is also thought to have a functional implication in compensating the gain in enthalpy upon binding by the loss of conformational entropy (89). This effect, termed isothermal enthalpy/entropy compensation, effectively uncouples binding-specificity from binding-strength, which enables specific binding partners to associate and dissociate reversibly. Recently this effect has been demonstrated in polyuridine-tract recognition of the splicing factor U2AF (90), and is known be valid for DNA-binding proteins (91). Interestingly, a recent study (92) revealed that Hfq65 and Hfq75 are defective in DNA binding, implicating the C-terminus as well as the distal surface in DNA interactions.

Binding of a poly(A) tract to six tripartite binding motifs at the distal face of HfqEc has been demonstrated in structural studies (23). However, as HfqEc65 was deficient in mediating low-temperature translational activation of *rpoS* mRNA by the sRNA DsrA (18) and HfqEc65 did not bind to *rpoS* mRNA containing the A-rich motifs implicated in distal site binding (30), these interactions are apparently not sufficient for the function of HfqEc in *rpoS* regulation. In addition, and in agreement with the observed annealing deficiency of HfqEc75, HfqSa and HfqEc (Figure 1A), the three proteins were defective in complementing full-length HfqEc in low temperature activation of *rpoS* [Figure 1B; (18)]. Furthermore, HfqSa and HfqEc did not complement HfqEc in RyhB-mediated

**SRCD spectroscopy reveals an interaction of the C-terminus with RNA**

In order to monitor changes in the structure of the C-terminal extension of HfqEc in the presence of RNA, we next used SRCD spectroscopy. Both HfqEc65 and full-length HfqEc bind the sRNA DsrA (82) and also the Hfq binding fragment of DsrA (DsrA34; data not shown) with comparable affinity, indicating that the C-terminus is not required for binding of the sRNA. As the SRCD spectra of HfqEc alone and in the presence of DsrA34 were indistinguishable (Figure 7), the known binding mode of DsrA to the inner core of the HfqEc hexamer (29), apparently does not affect the structure of the protein. In contrast to sRNAs, longer (m)RNA fragments did not bind to the HfqEc65 but depended on the presence of the C-terminus (18). Interestingly a 126 nt long fragment of *hfq* mRNA, which failed to bind to HfqEc65 but bound to full-length HfqEc (18), produced significant increases in the 185–200 nm region only upon binding to full-length HfqEc, reflecting an increase in the amount of ordered (\( \beta \)-strand or \( \alpha \)-helical) conformations for this part of Hfq. The increase is consistent with a combination of decrease of disordered and increase in ordered residues. Altogether this suggests that RNA binding leads to an ordering of the C-terminal extension of Hfq. Many intrinsically disordered proteins undergo transitions to more ordered states or fold into stable secondary or tertiary structures upon binding to their targets. They undergo coupled folding and binding, forming complexes with high specificity and relatively low affinity, which is critical for processes in which not only specific association but also subsequent dissociation of binding partners is required (87).
translational repression of sodB mRNA (16). When compared to self-annealing HfqSa reduced the annealing rate of the two oligonucleotides. Although we have not further studied this behaviour of HfqSa it appears possible that this protein binds one or the other RNA too tightly, thereby reducing self-annealing. In contrast, HfqEc85 was functional in the annealing assay (Figure 1A) and stimulated rpoS activation by DsrA at low temperature. Thus, this study implicates the region between amino acids 75 and 85 as a minimal portion supporting the interactions with longer RNA fragments, whereas the core region, i.e. HfqEc65 is fully proficient in sRNA binding. Thus, the structurally flexible and disordered C-termini appear to provide a moiety involved in tethering of longer and structurally diverse RNA molecules, which could be followed by stable accommodation of the substrate at the distal site, i.e. at the polyA binding motifs as suggested for rpoS mRNA (23). Such a rearrangement would be in line with the concept of isothermal enthalpy/entropy compensation, as discussed above. While we see a certain spread in the conformations of the C-termini from our SAXS analysis, with the best fitting conformations extending laterally away from the hexameric core—the crystallization process apparently selects conformers with the C-termini on the proximal face, indicating this conformation to be one of the low energy states of the molecule. Using NMR, SAXS and small angle neutron scattering, we have obtained evidence that the DsrA34 fragment binds to the proximal region and extends into the solvent (M. B.-F., unpublished data). Thus, annealing of the sRNA with the mRNA may occur either during the accommodation of mRNA at the distal site or after the mRNA is bound at the distal site.

Although these biophysical studies predestine the C-termini of HfqEc for interactions with RNA, they are seemingly at variance with the observations that the Hfq orthologues of M. jannaschii (HfqMj; 71 amino acids) and Listeria monocytogenes (HfqLm; 77 amino acids) complemented HfqEc in rpoS translation (at least at 37°C), and RyhB-mediated decay of sodB mRNA (26,93). While HfqEc has relatively short predicted intrinsically disordered segments in the N-terminus (the first 5–6 amino acids), in HfqMj the first 13 N-terminal residues are predicted to be disordered, while the C-terminus is not (Supplementary Figure S2), which could indicate a role of the N-terminus in RNA binding. Nonetheless, HfqSa, being with 77 amino acids as long as HfqLm with also similar pattern of predicted unstructured regions, was in our hands defective in RyhB-mediated translational repression of sodB mRNA (18). Thus, the possibility exists that Hfq proteins of different origin interact differently with distinct substrates. In this regard it is notable that HfqEc65, although defective in the annealing assay and in supporting translation of rpoS (Figure 1) was fully functional in phage Qβ replication.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
The group of D. Svergun, particularly Weifeng Shang, Adam Round, Alexei Kikhney and Manfred Roessle, at the X33 beamline, DESY, EMBL-Hamburg, is gratefully acknowledged for help with SAXS data-collection. Greg J. R. Ewing is acknowledged for help with shell scripting, and Euripedes de Almeida Ribeiro for helpful discussions. Access to the CD1 beamline at ISA (the Institute for Storage Ring Facilities, Aarhus University, Denmark) is acknowledged under the EU Integrated Infrastructure Initiative, Integrated Activity on Synchrotron and Free Electron Laser Science (IA-SFS), contract number RII3-CT-2004-506008.
FUNDING
The work in the K. Dj.-C., U.B. and R.K laboratories was supported by grants F1722 (K. Dj.-C.), F1720 (to U.B) and F1719 (to R.K.) in the framework of the Special Research Program (SFB17) on ‘Modulators of RNA fate and function’ by the Austrian Science Fund. U.K. Biotechnology and Biological Research Council (BBSRC) project grant (to B.A.W.); SRCD beamtime grants (to B.A.W.). Funding for open access charge: Austrian National Science fund.

Conflict of interest statement. None declared.

REFERENCES


