A novel dimerization motif in the C-terminal domain of the *Thermus thermophilus* DEAD box helicase Hera confers substantial flexibility

Dagmar Klostermeier\(^1\)* and Markus G. Rudolph\(^2\)

\(^1\)Division of Biophysical Chemistry, Biozentrum, University of Basel, CH-4056 Basel, Switzerland and
\(^2\)Department of Molecular Structural Biology, University of Göttingen, D-37077 Göttingen, Germany

Received September 23, 2008; Revised October 30, 2008; Accepted November 8, 2008

ABSTRACT

DEAD box helicases are involved in nearly all aspects of RNA metabolism. They share a common helicase core, and may comprise additional domains that contribute to RNA binding. The *Thermus thermophilus* helicase Hera is the first dimeric DEAD box helicase. Crystal structures of Hera fragments reveal a bipartite C-terminal domain with a novel dimerization motif and an RNA-binding module. We provide a first glimpse on the additional RNA-binding module outside the Hera helicase core. The dimerization and RNA-binding domains are connected to the C-terminal RecA domain by a hinge region that confers exceptional flexibility onto the helicase, allowing for different juxtapositions of the RecA-domains in the dimer. Combination of the previously determined N-terminal Hera structure with the C-terminal Hera structures allows generation of a model for the entire Hera dimer, where two helicase cores can work in conjunction on large RNA substrates.

INTRODUCTION

Helicases couple the energy from ATP hydrolysis to structural rearrangements of their nucleic acid substrates. Based on conserved sequence motifs, helicases are divided into superfamilies SF1, SF2 and SF3 (1). RNA helicases of the DEAD box family belong to SF2 and are involved in virtually all processes of RNA metabolism. DEAD box helicases share a helicase core domain that contains all conserved helicase motifs (2). This core domain is subdivided into two RecA-domains (3,4) that are flexibly connected. Additional domains may contribute to substrate binding, confer binding specificity or mediate interactions with other proteins. The N-terminal domain contains many conserved motifs involved in nucleotide-binding, such as the Q-motif, the Walker A motif and the DEAD box. The C-terminal RecA-domain contains motifs IV–VI that are essential for RNA binding. RNA helicases undergo large conformational changes during their catalytic cycle (5). In the closed, ATP-bound state the N- and C-terminal RecA-domains associate to form a continuous binding site for RNA and allow motifs from both domains to form an interaction network.

Closed conformations of RNA helicases were trapped in crystalline form in few instances, such as the exon junction complex (6,7) and the *Drosophila melanogaster* Vasa RNA helicase (8), which were determined in complex with an ATP analogue and a single stranded RNA oligonucleotide. Information on the open helicase core conformations was gained from structures of eIF4A (9), an uncharacterized *Methanococcus jannaschii* DEAD (mjDeaD) box helicase (10), UAP56 (11,12), Dhhlp (13) and DDX3X (14). Since these structures comprise the helicase core only, the location of any appendage domains is currently unknown. Most of the helicase core structures show a monomer, but there is some evidence from analysis of crystal contacts that the N-terminal domains of Hera (15), mjDeaD (10) and UAP56 (11,12) may dimerize in solution.

In the heat resistant RNA-dependent ATPase Hera from *Thermus thermophilus*, the helicase core is followed by a C-terminal extension that mediates interaction with 23S ribosomal RNA fragments and RNase P RNA (16,17). Hera is the first DEAD box helicase that forms a stable dimer in the absence of ligands, which sets it apart from other RNA helicases and raises the question as to how dimerization might be linked to RNA binding and helicase function. N-terminal truncations mapped the dimerization domain to the C-terminal half of Hera (17). We determined three crystal structures of the C-terminal part of Hera from *T. thermophilus* (termed Hera\(_{208-419}\)) on the way to a model of the complete Hera helicase. The structures display a number of unique features. First, Hera forms a dimer in the crystal,
which was also confirmed in solution studies. Second, the first 50 residues of the C-terminal domain form the dimerization motif, which adopts a novel fold. It is loosely packed and imparts a high degree of flexibility on the Hera dimer. Comparison of the three independent structures shows that this flexibility allows for drastic changes in the juxtaposition of the two helicase cores within the dimer. Third, while the C-terminal 90 residues could not be built into the electron density, their localization and extent are clearly visible and provide a first glimpse on this additional RNA-binding site outside the Hera helicase core. Thus, the C-terminal extension of Hera is a bipartite structure of a dimerization and an RNA-binding module.

MATERIALS AND METHODS

Protein production, purification, size exclusion chromatography and crystallization

Hera was produced and purified as described (17) but failed to crystallize. Two C-terminally truncated constructs, Hera_208–510 and Hera_208–419, were then generated and purified using the same protocol as for authentic Hera. These constructs eluted as dimers from a calibrated S200 gel permeation column (GE Healthcare). Analytical size exclusion chromatography was performed on a S200 10/300 GL column (GE Healthcare) in 50 mM Tris/HCl, pH 7.5, 500 mM NaCl. Full-length Hera from T. thermophilus was mixed 1:2 with reservoir and incubated in a microbatch setup. Tetragonal crystals of Hera_208–510 grew at 4°C from reservoir containing 0.1 M MES/NaOH pH 6.5, 15–18% ethylene glycol, 0.3–0.5 M NaI, 3–6% PEG20000 and are of space group P42_121 with one molecule per asymmetric unit. Two orthorhombic crystal forms of Hera_208–419 grew from reservoir containing 0.1 M Tris/HCl pH 8.5, 0.2 M (NH4)2SO4, 23% PEG3350 and 0.1 M Tris/HCl pH 7.0, 0.2 M (NH4)2SO4, 25% PEG3350, 5–10% (w/v) glucose or sucrose. Both crystal forms belong to space group P2_12121 and contain two molecules per asymmetric unit. Further, details for purification, crystallization and phasing of the Hera fragments will be given elsewhere (18).

Data collection, structure determination and refinement

Data were collected at 100K and reduced with the HKL programs (19) or with XDS (20) and SADABS (Bruker). The structure was determined by MAD phasing using data from a tetragonal crystal of Hera_208–510 containing a single selenium site per asymmetric unit (18). Model building was initiated from the coordinates of PDB-ID 1hv8, which were manually fit into the electron density map from the MAD experiment after density modification and phase extension. After rebuilding, the tetragonal structure could not be satisfactorily refined, probably due to the anisotropy and disorder, which was observed for >20 datasets collected from tetragonal crystals. In some cases merohedral twinning was present with the true space group being P2_12121 (a ≈ b) and a dimer in the asymmetric unit (data not shown). Although MAD-phasing using merohedrally twinned data is possible (21), the dataset used here was not twinned. Electron density is present for the C-terminal RNA-binding domain (RBD; Hera residues 420–510), which enables placement of this domain (see Results section). The orthorhombic structures were determined using the tetragonal structure as the starting model for molecular replacement with PHASER (22). Models were built in COOT (23) and refined with BUSTER-TNT (24) or PHENIX (25). 5% of reflections were reserved for Rfree cross-validation in all structures (26). Statistics are summarized in Table 1. Surface complementarity coefficients (27) and solvent accessible surface areas were calculated with the programs SC and AREAIMOL, respectively, as implemented in CCP4 (28), using a 1.7 Å radius probe. Possible hydrogen bonds and van der Waals contacts were detected with CONTACSYM (29) using default parameters. Inter-helical angles were calculated using HELIXANG (28). Figures were created with Bobscript (30) and rendered with Raster3D (31), or PyMol (www.pymol.org).

RESULTS

Full-length Hera from T. thermophilus was purified in a nucleotide-free form that binds and hydrolyzes ATP and is an active RNA helicase (17). Crystallization of the full-length Hera has been reported several years ago (16), but was irreproducible in our hands. In a divide et impera approach we cloned the genes coding for N- and C-terminal fragments (residues 1–207, 208–510 and 208–419) to determine their crystal structures separately and prepare a model for the full-length Hera. The structure of the N-terminal domain in complex with AMP was determined earlier (15). Here, we report three significantly different structures of the C-terminal Hera fragment 208–419 in two orthorhombic crystal forms. A third, tetragonal crystal form of Hera_208–510 was used for Se-Met MAD phasing (see Methods section) but the model built into the experimental electron density maps could not be refined satisfactorily (18). However, this model was of sufficient quality for molecular replacement of the orthorhombic Hera_208–419 crystal forms. The asymmetric unit of both forms contains two Hera_208–419 molecules that form dimers. Dataset 1 of crystal form 1 has one of the two RecA domains disorder, which was omitted from the model. This structure is referred to as ‘partial’ throughout. Another dataset of crystal form 1 and data from crystal form 2 allowed tracing of complete Hera_208–419 dimers, called dimer I and dimer II, respectively. The structures were refined to resolutions of 2.8 Å (dimer I) and 2.3 Å (partial structure and dimer II). Data collection and refinement statistics are summarized in Table 1.

Architecture of the Hera monomer

Hera_208–419 folds into a RecA domain connected to a dimerization domain that in authentic Hera is followed by a C-terminal RBD. The RecA-domain comprises a
Table 1. Data collection, phasing and refinement statistics

<table>
<thead>
<tr>
<th>Dataset</th>
<th>3EAR—form 1, partial</th>
<th>3EAS—form 1, complete, asymmetric</th>
<th>3EAQ—form 2, complete, asymmetric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection resolution range, (Å)</td>
<td>46.0–2.3 (2.36–2.30)</td>
<td>46.1–2.8 (2.88–2.80)</td>
<td>46.9–2.3 (2.34–2.30)</td>
</tr>
<tr>
<td>100% criterion (Å)b</td>
<td>2.3</td>
<td>2.8</td>
<td>2.3</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁</td>
<td>P2₁2₁2₁</td>
<td>P2₁2₁2₁</td>
</tr>
<tr>
<td>Cell dimensions (Å)</td>
<td>a = 41.6, b = 67.7, c = 183.8</td>
<td>a = 46.1, b = 70.8, c = 181.2</td>
<td>a = 62.6, b = 70.8, c = 101.9</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>23 505 (1375)</td>
<td>15 368 (1089)</td>
<td>20 794 (772)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>6.3 (6.3)</td>
<td>6.3 (6.3)</td>
<td>12.8 (10.3)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.9 (94.8)</td>
<td>99.9 (100)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>Mosaicity (°)</td>
<td>0.45</td>
<td>0.32</td>
<td>0.28</td>
</tr>
<tr>
<td>Rsym (%)</td>
<td>4.8 (66.1)</td>
<td>7.0 (79.0)</td>
<td>10.6 (95.0)</td>
</tr>
<tr>
<td>Average</td>
<td>14.7 (1.9)</td>
<td>10.1 (1.4)</td>
<td>12.9 (1.7)</td>
</tr>
<tr>
<td>Refinement resolution range (Å)</td>
<td>45.9–2.3</td>
<td>45.3–2.8</td>
<td>46.9–2.3</td>
</tr>
<tr>
<td>Rfree %</td>
<td>23.1/25.5</td>
<td>22.8/31.4</td>
<td>21.8/26.9</td>
</tr>
<tr>
<td>Number of residues/waters</td>
<td>256/29</td>
<td>415/0</td>
<td>416/7</td>
</tr>
<tr>
<td>Coordinate error (Å)</td>
<td>0.34</td>
<td>0.58</td>
<td>0.29</td>
</tr>
<tr>
<td>r.m.s.d. bond/angles (Å, °)</td>
<td>0.006/0.92</td>
<td>0.008/1.15</td>
<td>0.007/1.07</td>
</tr>
<tr>
<td>Ramachandran plot (%)</td>
<td>93.2/6.8/0</td>
<td>92.7/7.3/0</td>
<td>92.4/7.3/0.3</td>
</tr>
</tbody>
</table>

*a*Values in parentheses correspond to the highest resolution shell.

*b*The 100% criterion was calculated using SFTOOLS and represents the resolution in Å of a 100% complete hypothetical data set with the same number of reflections as the measured data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_{sym} = \frac{1}{N} \sum_{i} \frac{I_{obs}(h) - I_{calc}(h)}{I_{calc}(h)} )</td>
<td>0.006/0.92</td>
</tr>
<tr>
<td>( R_{free} = \frac{1}{N_{free}} \sum_{i} \frac{I_{obs}(h) - I_{calc}(h)}{I_{calc}(h)} )</td>
<td>0.008/1.15</td>
</tr>
<tr>
<td>( R_{free} = \frac{N_{free}}{N} \sum_{i} \frac{I_{obs}(h) - I_{calc}(h)}{I_{calc}(h)} )</td>
<td>0.007/1.07</td>
</tr>
<tr>
<td>Based on maximum likelihood.</td>
<td></td>
</tr>
</tbody>
</table>

The RNase P motif in Hera

Sequence alignments have shown that the region encompassing residues 372–386 exhibits weak homology to the signature motif of the RNase P-protein component (16). A connection of Hera and RNase P with respect to RNA recognition via the RNase P motif would be exciting, and indeed we have recently shown RNase P RNA to be a specific substrate of Hera (17). However, comparison of the crystal structures of Hera_208–419 and the RNase P protein from *Thermotoga maritima* [PDB-ID 1nz0, (35)] clearly shows that both proteins are dissimilar. Although both RNase P motifs form \( \alpha \)-helices (though not over the entire sequence in *T. maritima* RNase P), the structural context is very different (Figure 1d). Hera is all \( \alpha \)-helical in its C-terminal region while the RNase P-protein adopts a mixed \( \alpha/\beta \)-fold. A common function might exist for the RNase P motifs in RNA binding despite the presence of conserved arginines [Arg378, Arg383 and Arg386 in Hera (16)]. However, a Hera mutant where these arginines are replaced by alanines is still able to bind RNA, rendering the RNase P motif dispensable for RNA binding (17).

entric seven-stranded parallel \( \beta \)-sheet flanked by two and four \( \alpha \)-helices on either side (Figure 1a and b). A DALI search (www2.ebi.ac.uk/dali/) for structural homologues results in >50 unique structures with a Z-score >10, among them the Bacillus subtilis DEAD box helicase YxiN (Z-score 23.9, r.m.s.d. 1.8 Å over 151 residues) and the M. jannaschii DEAD box helicase (Z-score 23.0, r.m.s.d. 1.7 Å over 149 residues), but also the more distantly related helicase domains in *D. melanogaster* Vasa (Z-score 20.1, r.m.s.d. 2.2 Å over 146 residues; (8), the human exon junction complex (Z-score 21.2, r.m.s.d. 2.1 Å over 154 residues; (6,7), and hepatitis C virus NS3 (Z-score 13.0, r.m.s.d. 2.5 Å over 116 residues; (32), for which structures of nucleic acid complexes exist. The conserved helicase motifs IV–VI in the RecA domain of Hera are located close in space and interact with each other. A key residue in this network is Arg335 of motif VI that bridges motifs IV and VI. Phe244 in motif IV interacts with Arg325 via \( \pi \)-cation interactions, while Asp296 of motif V both neutralizes the positive charge of Arg325 and aligns this side-chain for interaction with Phe244 by two hydrogen bonds (Figure 1c). The helicase motifs thus form a platform for the second RecA domain of the helicase core (33).

The RecA domain is followed by a \( \alpha \)-helical extension (residues 370–419) that folds into a single turn of a left-handed super-helix. A BLAST search for this region (www.ncbi.nlm.nih.gov/BLAST) revealed no significant sequence homology to any other protein. A DALI search for structural homologues using this domain as a query resulted in only one remotely similar structure with a Z-score of 3.8, where scores <2.5 are defined as structurally dissimilar. The topology of HP0242, a *Helicobacter pylori* protein of unknown function [PDB-IDs 2ouf; MCGS, unpublished and 2bo3; (34)], is similar to Hera, also forming a left-handed super-helix. However, although both molecules use this domain to form dimers, the differences in length and angles of the helix segments lead to a large r.m.s.d. value of 3.7 Å over 50 residues for these structures, which should therefore be considered unrelated (Figure 1d). As a result, we conclude that this dimerization motif constitutes a novel fold that is currently unique to Hera.
It thus appears that the sequence homology between Hera and RNase P is coincidental.

**The Hera dimer**

Full-length Hera forms a dimer in solution as judged by gel permeation chromatography. Hera constructs encompassing residues 1–510, 208–510, 208–419 and 370–510 are apparent dimers in solution (17), putting the dimerization domain within the region encompassing residues 370–419. To prove the dimeric nature of Hera in solution, gel permeation chromatography was performed on mixtures of Hera_1–510 (full-length, 56 kDa per monomer) and Hera_208–419 (24 kDa per monomer). Provided rapid monomer exchange, the appearance of a heterodimeric species with intermediate molecular weight is expected.

When Hera_1–510 and Hera_208–419 were mixed, only two peaks corresponding to the individual proteins were observed in the chromatogram, indicating slow (or no) monomer exchange (data not shown). After heating the mixture to 65°C for 10 min, however, a third peak corresponding to a heterodimer appeared (Figure 2a). The population of this species increased further when the incubation time was increased to 20 min, consistent with slow exchange kinetics even at elevated temperatures. Controls with the equally treated isolated proteins yielded only one peak at the expected molecular weight of dimers, confirming that the additional peak in the mixture is not an artifact due to incubation at 65°C. The dimeric nature of Hera is confirmed by the crystal structure of the partial dimer (Figure 2b), and further corroborated by the crystal structures of the complete dimers (Figure 2c and d).

Figure 1. Hera structure sequence relationship and architecture of the Hera_208–419 monomer. (a) Sequence of Hera_208–419 with secondary structure elements indicated on the top. The numbering of sequence and secondary structure elements corresponds to the full-length Hera. Conserved helicase motifs IV–VI are colored blue, magenta and green, respectively. The putative RNase P motif is colored in red. (b) Stereo ribbon diagram of the Hera_208–419 monomer. The secondary structure elements, helicase and RNase P motifs from (a) are indicated. (c) Close-up showing the interactions of helicase motifs IV–VI. Asp296 of motif V connects to Arg325 motif VI via two hydrogen bonds (dashed lines). Arg325 stacks on Phe244 of motif IV. The view is rotated by 180° around the y-axis compared to (b). (d) Comparison of the left-handed super-helices in Hera_208–419 (left) and the hypothetical H. pylori protein HP242 (right). The corresponding monomers are colored identically in yellow and transparent grey. (e) The RNase P motifs in Hera (left) and in the protein component of the T. maritima RNase P (PDB-ID 1nz0; right) are predominantly α-helical but located in a very different structural context. The structures are shown with their RNase P motifs aligned (red).
The overall fold of the Hera monomer is replicated in the Hera dimers. Both dimer structures adopt a \( \alpha \)-shape with the helicase motifs IV–VI facing towards a flat surface that constitutes part of the interface with the N-terminal RecA domain in the closed conformation (Figures 1b and 4). The arrangement of the second RecA domains and the dimer interfaces are very different when the dimer structures are compared to each other, and thus indicate pronounced flexibility of Hera also in solution. Whereas the overall structure of dimer II is almost symmetric, with a pseudo-twofold axis perpendicular to the dimerization domain (Figure 2c), dimer I is highly asymmetric (Figure 2d). Superposition of the first RecA domains for dimers I and II shows a large displacement of the second RecA domain. The two Hera conformations are characterized by a 34\(^\circ\) rotation, which places equivalent residues up to 25\( \AA \) apart from each other (Figure 2e and f). One cause for this flexibility in Hera is a hinge region (red in Figure 2e and f) formed by residues Pro366–Pro369, which allows for different directions of the adjacent helix \( \alpha \)15. Comparison of the three Hera_208–419 structures shows the \( \alpha \)15 inter-helical angles to change in steps of 12\(^\circ\) (inset in Figure 2f). The hinge allows the C-terminal ends of helix \( \alpha \)15 to vary by up to 9\( \AA \) (at Ala385), with a concomitant amplification in the displacement of the RecA domain of the second monomer. This mobility can be further augmented by equivalent hinge movements in the second monomer. The hinge around Pro366–Pro369 is, however, not the sole origin of flexibility in the Hera dimer, as the dimer interface also shows considerable flexibility.

### The dimerization domain and flexibility in the Hera dimer

Two left-handed helical regions formed by Hera residues 370–419 associate to form a pseudo-symmetric dimer that has well-defined electron density (Figure 3a). No other
regions seem to be contributing to Hera dimerization, in accord with the solution studies. The predominantly hydrophobic dimer interface is both extensive and of high complementarity. While the partial structure buries a surface of 3300 Å², more than 5300 Å² are solvent inaccessible in dimer I and dimer II (Table 2). The average surface complementarity coefficient $S_c$ is $\sim$0.75, indicating a very good fit of the monomers, which is also mirrored in the large number of inter-subunit contacts. The interface is dominated by van der Waals interactions: around 200 are present, many of them centered at residues Trp377, Tyr392, Tyr395, Phe398, and helix $\alpha_{18}$ formed by residues Val408-Leu419 (see Supplementary Table 1 for a complete listing). The parallel orientation of helices $\alpha_{18}$ forms the hydrophobic core of the dimer interface (Figure 3b) and at first sight resembles a leucine zipper. The inter-helical angle in the prototypic GCN4 leucine zipper is $\sim$24°. In contrast, the $\alpha_{18}$ inter-helical angles in the partial structure, dimer I and dimer II are 52°, 54° and 58°, respectively, excluding a leucine-zipper type interaction. The r.m.s. differences in the $\alpha_{18}$ inter-helical angles also point to considerable flexibility in the dimer interface. There are a few polar contacts in the interface, most notably between the side-chains of Glu374/Arg401 and Tyr392/Glu409 (Figure 3a), but these are not present in all complexes. Indeed, a closer inspection of the interface contacts shows strong differences, both when comparing the different structures among each other and when comparing reciprocal interactions of the chains within one dimer. While the total number of contacts and buried surface areas in dimers I and II are roughly the same, there are fewer contacts and less buried surface in the partial structure, although the dimerization domain is complete. This discrepancy points to a less compact packing in the partial structure, indicating substantial plasticity in how the interface is organized.

Superposition of the dimerization domains onto helix $\alpha_{15}$ reveals that this domain can rotate about an axis parallel to the principal axis of helix $\alpha_{15}$. The extent of the rotation can be as large as 17°, resulting in maximum displacements of the apical atoms of the domain by up to 7 Å (Cα atoms of Ala404; Figure 3c and d). In conclusion, the overall structural plasticity of the Hera dimer is the sum of hinge movements and rigid body rotations of segments in the dimer interface. Due to this flexibility monomeric Hera may exist at elevated temperatures, but the observed slow monomer exchange underscores a high stability of the dimer even at physiologic temperatures for T. thermophilus.

**Construction of the Hera helicase core from its N- and C-terminal domains**

The Vasa helicase was crystallized in a closed conformation of its two RecA-domains in complex with an ATP analogue and a U$_{10}$ ssRNA (8), and represents one extreme of the conformational space accessible to

<table>
<thead>
<tr>
<th>Structure</th>
<th>Partial</th>
<th>Dimer I</th>
<th>Dimer II</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSA (Å²)$^a$</td>
<td>3355</td>
<td>5330</td>
<td>5361</td>
</tr>
<tr>
<td>$S_c$$^b$</td>
<td>0.736</td>
<td>0.771</td>
<td>0.751</td>
</tr>
<tr>
<td>Number of H-bonds/salt bridges$^c$</td>
<td>0/4</td>
<td>1/4</td>
<td>1/2</td>
</tr>
<tr>
<td>Number of vDW$^c$</td>
<td>199</td>
<td>221</td>
<td>228</td>
</tr>
<tr>
<td>Angle between $\alpha_{18}$ (°)</td>
<td>54</td>
<td>52</td>
<td>58</td>
</tr>
</tbody>
</table>

$^a$Total buried surface area (BSA) in the dimer calculated with a probe radius of 1.7 Å.

$^b$A surface complementarity coefficient of one would denote perfect complementarity.

$^c$Number of hydrogen bonds, salt bridges and van der Waals interactions in the dimer.
DEAD box helicases. The question now arises, how the structures of the dimeric Hera_208–419 and the N-terminal Hera domain would compare to Vasa and if a complete model for the Hera helicase core can be constructed without violating stereochemistry. The RecA-domains of Hera and Vasa superimpose closely with r.m.s.d. values of ~1.5 Å over 196 (N-terminal) and 125 residues (C-terminal), respectively. Indeed, Vasa-based alignment of the Hera dimer I and dimer II structures results in two feasible models for the complete Hera helicase core including the dimerization domain (Figure 4a and b). There are no stereochemical hindrances between the modeled RecA-domains that could not be rectified by slight conformational adjustments of loops and side-chains. Inspection of the RNA-binding site in Vasa shows that many of the side-chains contacting RNA have their Hera analogues in close proximity. Thus, not only is the construction of the artificial Hera helicase core stereochemically feasible, but it also makes biochemical sense and allows drawing conclusions on RNA binding by Hera (see below). Very similar models can be generated using the hepatitis C virus NS3 helicase (32) and the human exon junction complex (6,7) as templates. A possible limitation of these Hera helicase core models is the missing C-terminal domain. These additional modules are implicated in specific RNA binding (16,17,36–40) or in

Figure 4. Model for the dimeric Hera helicase and location of the RBD. (a, b) Construction of the Hera_1–419 dimers by superposition of the isolated Hera_N (PDB-ID 1gxs) and Hera_208–419 structures onto the Vasa-RNA complex. The view and color code for Hera_208–419 are the same as in Figure 2. Hera_N is colored orange, and the template Vasa is colored in white. (c) Stereo view of the location of the Hera RBD (residues 420–510) in crystals of Hera_208–510. Electron density from the MAD-phased tetragonal crystals is contoured at the 1σ level. This density is not explained by residues 208–419 and must thus belong to residues 420–510. The approximate volume for these residues is drawn as a transparent hull. The Hera_208–419 monomer is taken from dimer II and shown in gray. The connection between the Hera_208–419 C-terminus and the RBD is shown as a dashed line. (d) Electrostatic potential calculated for dimer II showing a distinct positively polarized patch (arrow) close to the RBD (location shown as a green surface). (e) View rotated 90° about the x-axis to emphasize the close proximity of the RNA-binding sites. The nucleotide AMP was taken from the Hera_N structure, the RNA from Vasa. Both are shown as stick models.
contributing to high affinity RNA binding by providing unspecific interactions with RNA (41–44). Their relative orientation with respect to the helicase core is currently unknown.

Localization of the Hera C-terminal RBD

The crystal structures presented here lack the C-terminal 90 residues that presumably form a separate domain, which was shown to be important for substrate RNA binding (17). The first crystals obtained and used for MAD-phasing were grown using a Hera construct encompassing residues 208–510, and thus included the C-terminal domain. Despite the fact that the experimental electron density generated from these crystals was of sufficient quality to trace residues 208–419 and also to use this model for successful molecular replacement of the orthorhombic crystal forms, no satisfactory refinement was obtained. What is obvious, however, from the electron density, is the location of the Hera C-terminal RBD. While we cannot assign secondary structure elements for the RBD, its envelope and location with respect to the remainder of the helicase are very clear. After building of the RecA-and dimerization domains, a globular volume of electron density wedged between these domains remains that must belong to the RBD (Figure 4c). The location of this domain is sensible as the C-terminus of Hera_208–419 points toward the un-modelled density. Leu419 is followed by a low complexity linker sequence (GGAPA) that can traverse the upper part of the RecA-domain towards the region of un-modelled electron density (dashed in Figure 4c).

Using the information of the localization of the RBD a complete model for the full-length dimeric Hera helicase can be assembled (Figure 4d and e). In this model, the RNA-binding sites of the dimers face each other, suggesting a possible coordinated RNA unwinding capability. Calculation of the electrostatic surface potential for the almost symmetric dimer II reveals a positively polarized patch generated by arginines 223, 345, 347 and 348 (arrow in Figure 4d and e). Most helicase cores do not display sequence specificity and, in agreement, the Vasa structure shows many protein–RNA phosphate backbone interactions. The positively polarized region around arginines 223, 345, 347 and 348 may serve as an unspecific binding site for the phosphate backbone of longer RNA substrates. In addition, this electropositive patch is uniquely positioned just underneath the location of the C-terminal RNA-binding domain, which further strengthens the notion that RNA substrates extend to the C-terminal domain via this path.

DISCUSSION

The crystal structure of the Hera_208–419 fragment confirms the location and interactions of the helicase motifs IV–VI involving residues Phe244, Asp296 and Arg325, respectively. The phenylalanine in motif IV was suggested as an anchor for the rigidity of the C-terminal RecA domain of the helicase core by π-cation interactions with an arginine in motif VI. Mutation of the motif IV phenylalanine to leucine or alanine in the SF2 helicase Ded1 results in loss of cooperative ATP/RNA binding (33). In Hera, Asp296 in motif V orients Arg335 by hydrogen bonding for optimal contact with Phe244, putting Arg335 at the interface of motifs V and VI. The same situation is observed in the mjDeaD helicase (10). A multiple sequence alignment of SF2 helicases [PFAM PF00271; pfam.sanger.ac.uk; (9)] shows that the Phe/Arg/Asp residue combination is not mandatory for SF2 helicase function. Instead, the network of interactions between these motifs can be established in the context of very different sequences or conformations with the phenylalanine replaced by hydrophobic and the aspartate replaced by small, often polar residues.

Possible binding of large RNA molecules to Hera

While the helicase core interacts non-specifically with RNA via residues in motif IV that contact the RNA backbone, the C-terminal domains in a few SF2 helicases mediate sequence-specific RNA binding. One well-studied example is the B. subtilis helicase YxiN, which contains a C-terminal domain of ~80 residues that confers specificity for hairpin 92 of the 23S rRNA (38). This domain folds into an RNA recognition motif that is found in many eukaryotic RNA-binding proteins, but employs a different RNA-binding mode (40). Although the Hera and YxiN RBDs have similar dimensions, they display no sequence homology and most likely adopt different structures, consistent with the impossibility to fit this domain into the electron density of the Hera RBD. Nevertheless, 23S rRNA fragments comprising hairpin 92 have also been identified as specific substrates for Hera, and binding of these RNAs requires the Hera RBD (17). A 32/9mer derived from the 23S rRNA, consisting of hairpin 92 and an adjacent 9-bp helix, is unwound by Hera in an ATP-dependent manner. The Hera model suggests a binding mode in which sequence- or structure-specific RNA binding to the Hera-RBD would serve as an anchor for hairpin structures, bringing the helicase core of Hera into proximity of RNA helices to be unwound. The electropositive patch generated by arginines 223, 345, 347 and 348 is located between the RBD and the helicase core, and could thus contribute to RNA positioning. Such a positioning would be even more important to direct unwinding or remodeling of larger RNA molecules that display a higher degree of conformational freedom. For example, the 379 nucleotide RNase P RNA is a specific Hera substrate and induces the closed conformation of the helicase core. The interaction of Hera with RNase P RNA also requires the RBD (17). The inherent flexibility of the RBD with respect to the dimerization motif, and of the dimerization motif with respect to the C-terminal RecA domain, would afford positioning of RNA with different sizes and structures for unwinding by the helicase core. Such a function is reminiscent of the suggested role of the basic CTDs of the DEAD box helicases Mss116p and CYT-19 (41,44,45). Clearly, a detailed description of the RNA-binding mode in Hera will have to await the structure of the Hera/RNA complex.
Implications of the RNA helicase dimer for RNA unwinding

Helicases in general display a variety of oligomerization states, from ring-shaped hexamers to dimers and monomers. Most DEAD-box helicases, e.g. the HCV-NS3 helicase (in the absence of the protease domain), function as monomers (46), but an enhanced activity and increased processivity of the dimeric form has been demonstrated (47,48). These observations point towards the existence of a functional dimer, whereas structural studies of NS3 support a monomer (32,49). The DEAD box helicases YxiN and DbpA exist as monomers in solution (50,51).

In contrast, the structures of Hera_208–419 from T. thermophilus described here represent first examples for a dedicated dimerization module in a DEAD-box RNA helicase. Based on the large buried surface area and high surface complementarity and the slow monomer exchange even at high temperatures it is likely that Hera forms a stable dimer that is not meant to dissociate during the catalytic cycle. Dimerization is a common strategy for protein stabilization in organisms living at high ambient temperatures (52). However, Hera homologues from the closely related mesophilic Deinococcus geothermalis (Q1J0S9) and the moderately thermophilic D. radiodurans (DR_1624) contain sequences with similarities to the Hera dimerization motif (17), arguing against a purely stabilizing role and in favor of a functional role of the helicase dimer.

The high structural similarity of the Hera N- and C-terminal RecA-domains compared to the Vasa and eIF4A-III helicases enabled us to construct a model for the closed Hera dimer. Inclusion of the information for the position of the RBDS results in an impressive architecture without any severe stereochemical hindrance (Figure 4d and e). The RNA-binding sites of the helicase cores face towards each other, suggesting that the two subunits could remodel RNA by simultaneously acting on the same RNA molecule, possibly in a coordinated fashion, which may be beneficial for efficient remodeling of large RNA substrates. To answer this question, mechanistic studies of monomeric Hera and structural information of the authentic Hera dimer bound to a large RNA would be essential.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

We thank I. Hertel, K. Gasow and A. Schmidt for excellent technical assistance and M. Linden for purifying full-length Hera. We also thank the colleagues at the SLS and BESSY synchrotrons for beamtime and guidance, and G. Sheldrick for providing a conversion program from XDS to SADABS reflection data format.

FUNDING

Deutsche Forschungsgemeinschaft (to M.G.R.); VolkswagenStiftung and the Swiss National Science Foundation (to D.K.). Funding for open access charge: Swiss National Science Foundation.

Conflict of interest statement. None declared.

REFERENCES


