Nucleic Acids Research, 1997, Vol. 25, No. 14 **2737–2744**

A biologically active 53 kDa fragment of overproduced alanyl-tRNA synthetase from *Thermus thermophilus* HB8 specifically interacts with $tRNA^{Ala}$ acceptor helix

Armin Lechler, Andreas Martin, Tilman Zuleeg, Stefan Limmer and Roland Kreutzer*

Laboratorium für Biochemie, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

Received April 22, 1997; Revised and Accepted May 28, 1997

DDBJ/EMBL/Genbank accession no. Y08363

ABSTRACT

The *alaS* gene encoding the alanyl-tRNA synthetase (AlaRS) from *Thermus thermophilus* HB8 was cloned and sequenced. The gene comprises 2646 bp, corresponding to 882 amino acids, 45% of which are identical to the enzyme from *Escherichia coli*. The *T.thermophilus* AlaRS was overproduced in *E.coli*, purified and characterized. It has high thermal stability up to ~65°C, with a temperature optimum of aminoacylation activity at ~60°C, and will be valuable for crystallization. The purified enzyme appears as a dimer with a specific activity of 220 U/mg and k_{cat}/K_M values of 118 000/s/M for alanine and 114 000/s/M for ATP. By genetic engineering a 53 kDa fragment of AlaRS comprising the N-terminal 470 amino acids (AlaN470) was also overproduced and purified. It is as stable as entire AlaRS and sufficient for specific aminoacylation of intact $tRNA^{Ala}$, as well as acceptor stem microhelices with a G3–U70, but not U3–A70, I3–U70 or C3–U70, base pair. The reduced binding strength of such microhelices to AlaN470 enabled, due to the resulting fast exchange of the microhelices between free and complexed states, preliminary NMR analyses of the binding mode and intermolecular recognition.

INTRODUCTION

Alanyl-tRNA synthetase (AlaRS), which belongs to class II of the aminoacyl-tRNA synthetases (aaRSs; 1), is one of the few aaRSs whose three-dimensional structure has not been solved to date. Since the known aaRSs are composed of several different domain modules in various combinations (2), a knowledge of each individual aaRS contributes to our understanding of the evolution of the genetic code. Also, the modules used for tRNA binding and recognition and their modes of interaction are surprisingly varied. In this respect AlaRS is particularly interesting, since the organization of this enzyme in domains and their functions are not yet fully understood.

Several experimental approaches have been used to identify common structural features of AlaRSs from different organisms and to determine the functions of the various enzyme regions (see 3 for a review). The N-terminal 368 amino acids of *Escherichia coli* AlaRS comprise the active site domain, where the three sequence motifs characteristic of class II aaRSs are found. This module is sufficient for alanylation (4). It is believed to interact with the G1–C72 base pair of $tRNA^{Ala}$ by approaching the acceptor stem from the major groove side (5). A retroviral-like zinc binding motif situated between motifs 2 and 3 contributes to tRNA recognition by this module (6,7). An additional module (amino acids 385–461), however, is needed for specific interaction of this enzyme with $tRNA^{Ala}$ by minor groove interactions at its conserved acceptor helix base pair G3–U70 (8). This base pair is the major determinant for the identity of $tRNA^{Ala}$. Amino acids 699–808 resemble a module which is believed to be responsible for oligomerization of *E.coli* AlaRS. This domain is different in the monomeric eukaryotic enzymes. At least two other modules with unknown function are found in the known AlaRSs.

Truncated tRNA model molecules, consisting of the acceptor stem of $tRNA^{Ala}$ from *E.coli*, are recognized by AlaRS as substrates and efficiently aminoacylated with alanine, provided the proper base pair is present at the critical 3–70 position (9). It has become a crucial question whether a helical distortion produced by the G3–U70 base pair (10) or a functional group on it (11) is required for recognition of the RNA by AlaRS. In addition to elucidation of the three-dimensional structure of the AlaRS–$tRNA^{Ala}$ complex by X-ray crystallographic analysis, NMR studies in particular are promising to clarify the mode of interaction with this acceptor helix base pair.

In this work we report cloning, sequence determination and overexpression of the *alaS* gene from *Thermus thermophilus* and purification of AlaRS and its biologically active monomeric fragment. The sequence information and the availability of large amounts of this thermostable enzyme are valuable for further structural investigations, since crystallization of AlaRS from mesophilic organisms has not been successful to date, for unknown reasons. In preliminary NMR experiments the N-terminal 53 kDa fragment of AlaRS (AlaN470) has proved particularly useful for studying its interactions with the acceptor stem of $tRNA^{Ala}$ microhelices.

*To whom correspondence should be addressed. Tel: +49 921 552433; Fax: +49 921 552432; Email: roland.kreutzer@uni-bayreuth.de

MATERIALS AND METHODS

Bacterial strains and plasmids

Thermus thermophilus HB8 (ATCC 27634) was grown as previously reported (12). *Escherichia coli* strains DH10B^{TM} (Gibco BRL, Bethesda, MD) and BL21(DE3)pLys (Novagen, Germany; 13) were grown in Luria Bertani (LB) nutrient medium (14) or Terrific broth (TB; 15). For cloning and sequencing plasmid pUC18 (16) and for expression plasmid pET-28c (Novagen, Germany; 13) were used.

Purification of AlaRS from *T.thermophilus*

The first purification steps of AlaRS from *T.thermophilus* involving anion exchange chromatography on Q-Sepharose-FF and gel permeation chromatography on Sephacryl S200 HR were performed essentially as previously described for purification of overproduced elongation factor Tu from *T.thermophilus* (17). Fractions with AlaRS activity were pooled and the protein precipitated by addition of solid ammonium sulfate to 70% saturation. The precipitate was dissolved in a minimal volume of buffer P10 (10 mM potassium phosphate, pH 7.0). The solution was dialyzed against 2× 2 l buffer P10, centrifuged and the particle-free supernatant loaded on a hydroxyapatite column (20 × 1 cm, particle size 15 µm, flow rate 1 ml/min; Merck, Germany) equilibrated with buffer P10. After the flow-through was eluted, the column was developed with a linear gradient of 10–200 mM potassium phosphate in 120 ml buffer P10. The fractions containing AlaRS were pooled, precipitated as above, dissolved in a minimal volume of buffer A [50 mM Tris–HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM 2-mercaptoethanol, 20 µM PMSF and 5% (v/v) glycerol] and dialyzed against 2× 1 l buffer A. The supernatant after centrifugation was loaded on a heparin–Sepharose CL-6B column (10 × 1 cm, flow rate 1 ml/min; Pharmacia Biotech, Freiburg, Germany). The column was developed with a linear gradient of 0–125 mM KCl in 100 ml buffer A. The fractions containing AlaRS were pooled and the protein precipitated as above. For use, the AlaRS precipitate was dissolved in an appropriate buffer and dialyzed. As assessed by SDS–PAGE, AlaRS was ~95% pure.

Purification of AlaRS from overproducing *E.coli*

The overproducing bacteria, *E.coli* BL21(DE3)pLys with one of the T7 promoter plasmids, were grown at 30°C in LB or TB in the presence of 40 µg/ml kanamycin and 25 µg/ml chloramphenicol. Fermentation of bacteria and the first purifications steps of AlaRS or AlaN470, involving heat treatment at 65°C and anion exchange chromatography on Q-Sepharose-FF, were performed essentially as previously described for purification of overproduced phenylalanyl-tRNA synthetase from *T.thermophilus* (18). AlaRS and AlaN470 was selectively precipitated with 45 and 50% saturated $(NH_4)_2SO_4$, respectively, and loaded on a hydroxyapatatite column as described. Collected fractions with AlaRS activity were dissolved in buffer A. Imidazole was added to a final concentration of 5 mM. A Ni-HiTrap® Chelating affinity column (volume 5 ml, flow rate 1 ml/min; Pharmacia Biotech, Freiburg, Germany) was equilibrated with buffer AI (buffer A + 5 mM imidazole) and saturated by addition of 2.5 ml 100 mM $NiSO_4$. The protein solution was loaded on this column. After elution of the flow-through, the column was developed with a linear gradient of 5–300 mM imidazole in 50 ml buffer AI. The fractions containing apparently homogeneous AlaRS or AlaN470 were pooled, dialyzed against 2× 1 l buffer A and finally against 100 ml buffer A + 50% glycerol and stored at –20°C.

Cyanogen bromide cleavage and protein sequencing

For cyanogen bromide cleavage 80 µg protein were dialyzed against water and dried in a vacuum centrifuge. An estimated 100-fold molar excess of cyanogen bromide dissolved in 70% (v/v) trifluoroacetic acid over AlaRS methionines was added and incubated at room temperature for 12 h. After adding 0.5 ml water the peptide solution was lyophilized; this was repeated twice. The peptides were separated by SDS–PAGE and blotted onto polyvinylidene difluoride (PVDF) membrane (Schleicher & Schüll GmbH, Dassel, Germany). The N-terminal amino acid sequences were determined with an ABI 473A protein sequenator as previously described (19).

Recombinant DNA technology

Standard recombinant DNA methods were performed according to the protocols of Ausubel *et al*. (20) and Sambrook *et al*. (15). PCR was carried out in 2 mM $MgCl_2$, 200 µM dNTPs,100 µM primers, 2 µg *T.thermophilus* genomic DNA partially digested with *Sau*IIIA as template, 2.5 U Taq DNA polymerase (Fermentas Ltd, Vilnius, Lithuania) in 10 mM Tris–HCl, pH 8.8, 50 mM KCl under the following conditions (in a volume of 100 µl): 5 min at 95°C; 30 cycles of 1 min at 95°C, 50 s at 40°C and 1 min at 72°C; 7 min at 72°C.

Plasmid constructions

The determined amino acid sequences (data not shown) were used to design degenerate oligodeoxyribonucleotides taking into account the peculiar codon usage of *T.thermophilus*. Oligodeoxyribonucleotides Ala 1 [ATGAAGCG(CG)AC(CG)GC(CG)GAGATCCG(CG)GAGAAGTT] and Ala 2 [CCCTTCTTGAAGATCTC(CG)CGCTCGAA(CG)GCCAT] were successfully used in PCR to amplify a product of 1.3 kb. This fragment was used as a probe to identify a 9 kbp *Sph*I fragment in a partial genomic DNA library constructed in pUC18. The corresponding plasmids were designated pUCAlaRS Sph9-1 and pUCAlaRS Sph 9-3 (insert in opposite orientation). The *alaS* gene was further subcloned, resulting in plasmids pUCAlaRS 4H1 (pUC18 with a 4 kbp *Sph*I–*Hin*dIII fragment comprising the 3′-portion of the *alaS* gene) and pUCAlaRS 3.4H1 (pUC18 with a 3.4 kbp *Hin*dIII fragment comprising the 5′-portion of the *alaS* gene), which were used to determine the nucleotide sequence of the *alaS* gene region.

For overexpression of AlaRS and AlaN470 the vector pET-28c, which allows purification of N-terminally histidine-tagged fusion proteins (20 additional amino acids at the N-terminus MGSSHHHHHHSSGLVPRGSH), was used. A *Nde*I restriction site overlapping the start codon of *alaS* was introduced by PCR into a 415 bp DNA fragment comprising nucleotides 1–404 of the *alaS* gene using primers AlaRSNde (TAATCAAACATATGCGCACGGCGGAGATCC) and Ala P2 (GCCTCGTCGTCGTCCTCAAAGACCGTGACC). The amplification product was hydrolyzed with *Nde*I and *Hin*dIII (internal restriction site) and ligated with expression vector pET-28c. A 4 kb *Hin*dIII fragment comprising the 3′-portion of the *alaS* gene was linked using the internal *Hin*dIII site. *Escherichia coli* BL21(DE3)pLys was transformed with the resulting expression plasmid (pETAlaRS1). To obtain an

expression plasmid for AlaN470, a stop codon (at AlaRS amino acid position 471) and an adjacent *Eco*RI restriction site were simultaneously introduced into the *alaS* gene by site-directed mutagenesis (21) using oligodeoxyribonucleotide Ala Eco/am (GCCTCGAGGGCGTTCTAGCCCAAGAATTCCGTGGCCC-CCC). The *Eco*RI site was used to identify the mutated clones. The changed nucleotide sequence in the resulting plasmid (pETAlaRS471am), as confirmed by sequencing, reads 5′-GCCA-CG**GAATTC**TTGGGC**TAG**AAC-3′ (*Eco*RI site and stop codon in bold). This plasmid expresses the 470 N-terminal amino acids of *T.thermophilus* AlaRS (AlaN470).

Activity test

The specific activities of AlaRS and AlaN470 were determined by a filter binding assay in 100 mM sodium cacodylate, pH 7.6, 10 mM KCl, 8 mM $MgSO_4$, 2 mM rATP, 5 mM 2-mercaptoethanol, 20 µM [^{14}C]alanine (120 Ci/mol; Amersham Buchler, Braunschweig, Germany), 70 µM *T.thermophilus* $tRNA^{bulk}$ (isolated according to the protocol of Zubay; 22) at 60°C in a volume of 60 µl. The protein concentration was determined by the BioRad (Hercules, CA) assay. One unit (U) of AlaRS activity was defined as the amount of enzyme needed to attach 1 nmol [^{14}C]alanine to the tRNA in 1 min at 60°C. For determination of the temperature dependence of the specific enzyme activity the sample was equilibrated at the respective temperature for 5 min before the reaction was started by addition of the enzyme. Prior to use the tRNA was preincubated at 60°C for 5 min and then added to the assay. To determine the aminoacylation of microhelices derived from the $tRNA^{Ala}$ acceptor stem, the chemically synthesized oligoribonucleotides (see below) were annealed by heating to 80°C for 3 min and slow cooling to room temperature. The resulting duplices at a concentration of 30 µM were assayed at 40°C with 30 nM enzyme in a total volume of 70 µl.

Size exclusion chromatography

The molecular masses of proteins were determined by size exclusion chromatography on Superdex™ 200 HR 10/30 (Pharmacia Biotech, Freiburg, Germany) by comparison with standard proteins (BioRad, Hercules, CA): thyroglobulin, 670 kDa; bovine γ-globulin, 158 kDa; chicken ovalbumin, 44 kDa; equine myoglobulin, 17 kDa; cyanocobalamin, 1.35 kDa. Aliquots of 100 µg purified protein were analyzed by developing the column with 50 mM Tris–HCl, pH 7.5, 10 mM $MgCl_2$, 150 mM KCl.

ATP/pyrophosphate exchange reaction

The reaction mixture (300 µl) contained 100 mM Tris–HCl, pH 7.5, 10 mM $MgCl_2$, 2 mM KF, 2 mM [^{32}P]pyrophosphate (0.5 mCi/mmol; DuPont NEN, Bad Homburg, Germany). ATP and alanine concentrations were in the range 40–600 and 100–800 µM respectively. After preincubation for 3 min at 50°C the reaction was started by addition of 10 nM enzyme. Aliquots of 40 µl were withdrawn at 2, 4, 6, 8 and 12 min and the reaction was stopped by addition of 70 µl 10% perchloric acid containing 300 mM sodium pyrophosphate. Aliquots of 100 µl of this solution were filtered through activated carbon paper (Schleicher & Schüll, Dassel, Germany). After washing with 5 ml water and 5 ml ethanol the filter was dried and the radioactivity of the absorbed ATP was counted. Kinetic data were analyzed using the Eadie–Hofstee and Lineweaver–Burk plots (23).

Thermal denaturation analysis

The protein solutions (1.5 and 3.7 µM for AlaRS and AlaN470, respectively) in 50 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 50 mM KCl were heated in a spectrophotometer (Hewlett Packard model 8452A) in 1°C steps with a linear temperature gradient from 20 to 90°C. The UV spectra in the range 240–350 nm were recorded after equilibration of the probe at each temperature for 3 min. Protein unfolding was determined by following the change in the absorbance difference at two wavelengths ($\Delta A_{274-286}$).

Chemical RNA synthesis, purification and NMR sample preparation

The oligoribonucleotides (7mer and 11mer strands) were chemically synthesized on a Gene Assembler Plus (Pharmacia Biotech, Freiburg, Germany) using the H-phosphonate method as described previously (24). RNA synthons were produced as specified (25). The oligoribonucleotides were purified by HPLC on a Nucleogen-DEAE 500-7 column (Macherey-Nagel, Germany).

For the NMR measurements the oligoribonucleotides (~0.25 mM strand concentration for one-dimensional studies, ~1.3 mM for the NOESY experiments) were dissolved in 0.5 ml H_2O/D_2O (9:1 v/v) with 100 mM NaCl, 4 mM $MgCl_2$, 10 mM sodium phosphate, pH 7. The strands were annealed by heating to 80°C for 3 min and slow cooling to room temperature. Approximately 0.2 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was added as an internal chemical shift reference.

NMR spectroscopy

All NMR spectra were acquired on a DRX 500 spectrometer (Bruker), equipped with an Aspect Station computer, at a proton resonance of 500 MHz. For the one-dimensional measurements the 1-3-3-1 pulse sequence (26) was used to suppress the strong water signal. The NOESY spectra were recorded using a 1-1 'jump–return' sequence (27) for the last (read) pulse with the excitation maximum in the imino region (~12.8 p.p.m.). Eight thousand data points in t_2 and 512 t_1 increments were used in the phase-sensitive mode, giving 4000 × 1000 two-dimensional NMR spectra after Fourier transformation. NOESY spectra were apodized by π/2-shifted squared sine functions in both dimensions. Data processing and spectra analysis were performed on Hewlett-Packard 9755 or INDIGO2 (Silicon Graphics) workstations using the NDEE program package developed by F.Herrmann (Symbiose Software, Bayreuth, Germany). The spectral width was 10 kHz, the mixing time 300 ms.

RESULTS

Cloning and sequencing of the *alaS* gene

To obtain sufficient amounts of AlaRS for crystallization and NMR studies via overexpression, as well as to enable mutagenesis, we first cloned the gene encoding AlaRS (*alaS*) from *T.thermophilus*. About 300 µg AlaRS were purified from *T.thermophilus* cells using standard chromatographic procedures. The enzyme was identified by its specific alanylation activity on $tRNA^{bulk}$ from *T.thermophilus*. To gain sequence information, the N-terminal amino acid sequences of the protein and of cyanogen bromide fragments were determined (data not shown). This was used to design gene probes, which were successfully used to clone the *alaS* gene on a 9 kb *Sph*I fragment. After subcloning the relevant

region, the nucleotide sequence of 4019 bp covering the *alaS* gene was determined on both strands and deposited in the EMBL database (accession no. Y08363). The *alaS* gene comprises 2646 bp, corresponding to 882 amino acids of the encoded AlaRS.

The protein is similar to the other known AlaRSs, e.g. 45% of the amino acids are identical in the corresponding enzyme of *E.coli* (Fig. 1). Sequence similarities are particularly pronounced not only around the three signature motifs of class II aaRSs but also around amino acids 400–411 (with respect to *T.thermophilus* AlaRS), as well as in a domain (*T.thermophilus* amino acids 500–716) of unknown function. The zinc binding motif reported for *E.coli* AlaRS (*E.coli* amino acids 179–192) was not found at a homologous position of *T.thermophilus* AlaRS (6). Cys666 (numbered including the N-terminal methionine) in the C-terminal region, which is important for aminoacylation activity of *E.coli* AlaRS (28), is also present in the *T.thermophilus* enzyme.

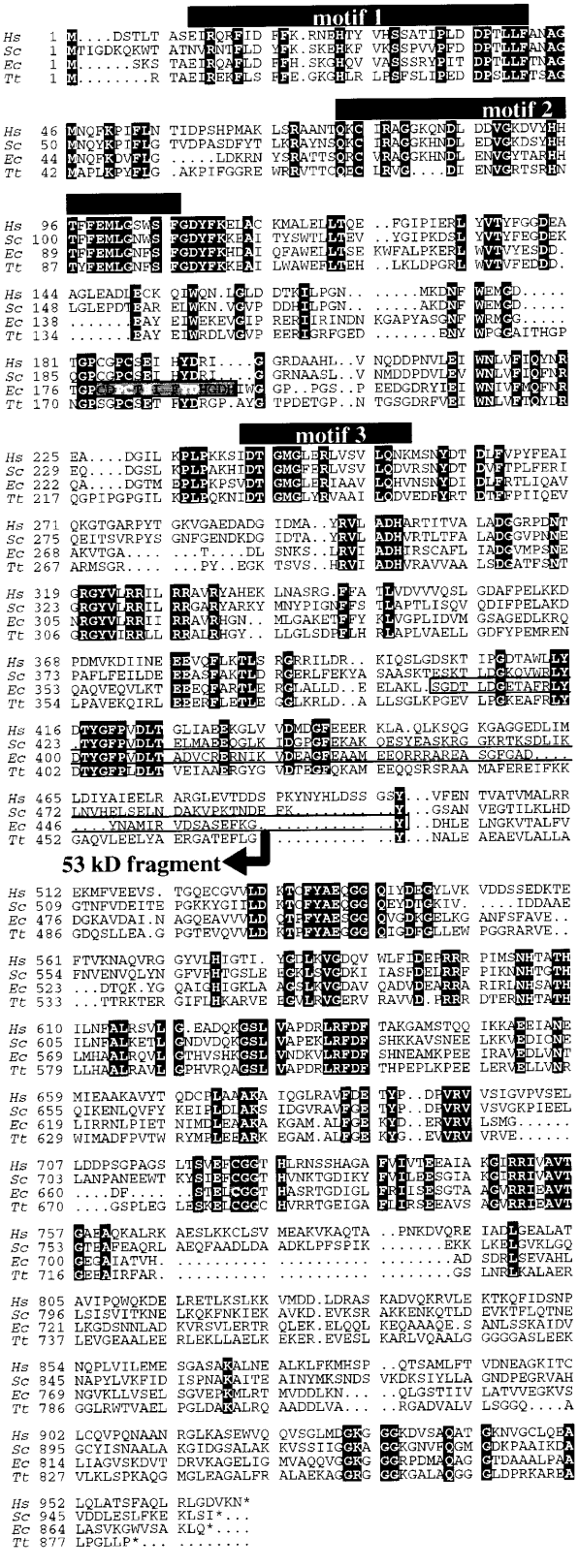

Figure 1. Alignment of the *T.thermophilus* AlaRS amino acid sequence (*Tt*) with those of *Homo sapiens* (*Hs*; EMBL accession no. D32050), *Saccharomyces cerevisiae* (*Sc*; U18672) and *E.coli* (*Ec*; J01581). In all four sequences identical amino acids are displayed on a black background. The three signature motifs of class II aaRSs are indicated by black bars above the sequences. The proposed retroviral-like metal binding motif of the *E.coli* AlaRS (6) is marked by a grey box, the region required for recognition of the G3–U70 base pair (8) is in an open box. The extent of the 53 kDa fragment of the *T.thermophilus* AlaRS (AlaN470) is indicated by an arrow and Cys666 in the *E.coli* AlaRS, which is important for aminoacylation (28), is underlayed in grey.

Overproduction in *E.coli* and purification of *T.thermophilus* AlaRS and its N-terminal 53 kDa fragment (AlaN470)

For overproduction of AlaRS from *T.thermophilus* and AlaN470 in *E.coli* we used pET-28c as the expression vector. This appended a histidine tag and a thrombin cleavage site at the N-terminus of the protein, thus facilitating purification. A *Nde*I restriction site overlapping the start codon of *alaS* was introduced by means of PCR, the *alaS* gene was put together from two neighboring restriction fragments and inserted into the vector (resulting in pETAlaRS1). For production of AlaN470, comprising the first 470 amino acids of AlaRS, codon 471 of the *alaS* gene was replaced by an amber stop codon by site-directed mutagenesis (pETAlaRS471am). Both the entire AlaRS and AlaN470 were overproduced in *E.coli* and purified by heat treatment of the crude cellular extract and chromatographies on Q-Sepharose, hydroxyapatite and immobilized nickel (Table 1). The proteins, apparently homogeneous on Coomassie Blue stained SDS gels (data not shown), were used for subsequent investigations.

Characterization of the overproduced enzymes

Thermal denaturation of AlaRS and AlaN470 was determined following the change in UV absorption upon heating of the protein solutions (Fig. 2). Both AlaRS and AlaN470 are stable up to ~65°C. Complete denaturation occurred at ~85°C, which was due to irreversible precipitation of the proteins. Slight differences in the denaturation kinetics can be attributed to a lack of cooperativity in monomeric AlaN470. Maximum aminoacylation was determined at ~60°C. It has, however, to be kept in mind that spontaneous deacylation of aminoacyl-tRNA at high temperatures may disguise a lower temperature optimum.

Table 1. Purification of the overproduced AlaRS and AlaN470 from 189 and 160 g respectively of *E.coli* BL21(DE3)/pLys + pET AlaRS (pET AlaRS471am)

	Protein (mg)	Total activity (U)	Specific activity (U/mg)	Purification factor	Yield (%)
AlaRS					
Crude extract	12 936	122 892	9.5		100
Crude extract after heat treatment (65°C)	3 620	111 858	30.9	3.3	91
Precipitation with $(NH_4)_2SO_4$ (45% saturation)	986	78 584	79.7	8.4	64
Q-Sepharose	311	45 095	145	15.3	37
Hydroxyapatite	156	25 740	165	17.4	21
Ni-HiTrap® Chelating	87	19 140	220	23.2	16
AlaN470					
Crude extract	11 500	8 625	0.8		100
Crude extract after heat treatment (65°C)	1 992	7 946	3.9	5.3	92
Precipitation with $(NH_4)_2SO_4$ (50% saturation)	715	5 999	8.4	11.2	69
Q-Sepharose	160	3 118	24.5	32.6	45
Hydroxyapatite	99	2 914	29.4	39.2	34
Ni-HiTrap® Chelating	68	2 078	30.6	40.8	24

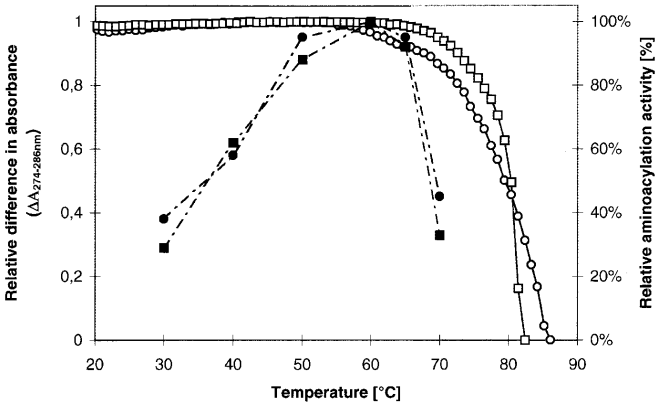

Figure 2. Thermal denaturation and temperature dependence of aminoacylation activities of AlaRS (□ and ■, respectively) and AlaN470 (○ and ●, respectively). Denaturation is followed by the relative difference in absorbance ($\Delta A_{274–286\,nm}$, left ordinate), where 1 means the lowest difference determined, corresponding to the native state of the protein. The temperature dependence is expressed as the relative aminoacylation activity (right ordinate), with 100% corresponding to the highest detectable specific enzyme activity (see Materials and Methods for details).

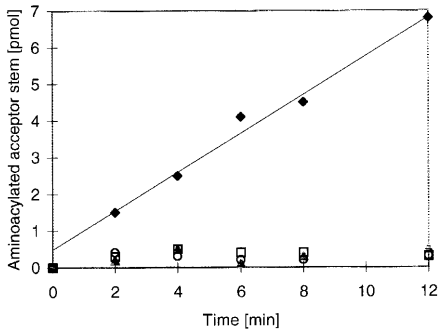

Figure 3. Aminoacylation activity at 40°C of AlaN470 (30 nM) using $tRNA^{Ala}$ acceptor stem helices as substrates (30 µM) with G3–U70 (◆), C3–A70 (▲), U3–A70 (□) and I3–U70 (○) base pairs. The values refer to the total activity in a 70 µl assay.

The molecular masses of AlaRS and AlaN470 were determined by size exclusion chromatography (data not shown). The experimentally determined mass of purified AlaRS (205 kDa) agrees fairly well with the two-fold mass calculated from the sequence (195 kDa), which clearly demonstrates that it is a dimer. In contrast, AlaN470 has an apparent mass of 48 kDa, corresponding to that calculated for the monomeric fragment (53 kDa).

The k_{cat}/K_M values for entire dimeric AlaRS determined by the ATP/pyrophosphate exchange reaction were 1.18×10^5/s/M for alanine and 1.14×10^5/s/M for ATP. The respective values for monomeric AlaN470 were 4.58×10^5 and 2.33×10^5/s/M. Specific activities for aminoacylation of $tRNA^{bulk}$ from *T.thermophilus* at 60°C were 220 and 31 U/mg for AlaRS dimer and for AlaN470, respectively.

AlaN470 specifically alanylated the acceptor stem microhelix with a G3–U70 base pair (see inset of Fig. 4; positions relative to entire $tRNA^{Ala}$), though the estimated catalytic efficiency may be at least two orders of magnitude lower compared with that with the native components (Fig. 3). However, no detectable aminoacylation activity of microhelices with U3–A70, I3–U70 or C3–A70 base pairs was found, indicating specific recognition of this critical wobble base pair even by the truncated enzyme, as reported for the corresponding molecule of *E.coli* (8).

Interaction of AlaN470 with $tRNA^{Ala}$ acceptor helix studied by NMR spectroscopy

In Figure 4a that part of the imino resonance region of the acceptor stem duplex is displayed which contains the signals due to base pairs 2 (G2–C71, ~12.51 p.p.m.) and 1 (G1–C72, 12.07 p.p.m.), the uracil of base 3 (G3–U70, 11.98 p.p.m.) and the guanine of the same wobble base pair (10.77 p.p.m.). The same spectral region is shown in Figure 4b after addition of ~3 mg AlaN470 (yielding a concentration of ~0.11 mM) to a 0.26 mM

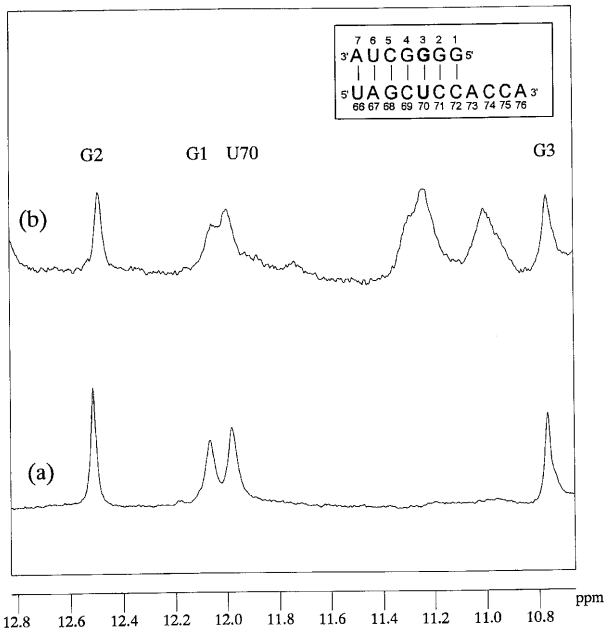

Figure 4. Part of the imino resonance region of the proton NMR spectra of the $tRNA^{Ala}$ acceptor stem duplex at a temperature of 310 K. (**a**) Acceptor duplex (0.26 mM) without protein; (**b**) acceptor duplex with ~3 mg AlaN470 (corresponding to a concentration of ~0.11 mM). The sample solutions contained 100 mM NaCl, 4 mM $MgCl_2$, 10 mM sodium phosphate, pH 7. The assignment is indicated in the upper trace. The inset displays the sequence of the acceptor stem duplex of $tRNA^{Ala}$.

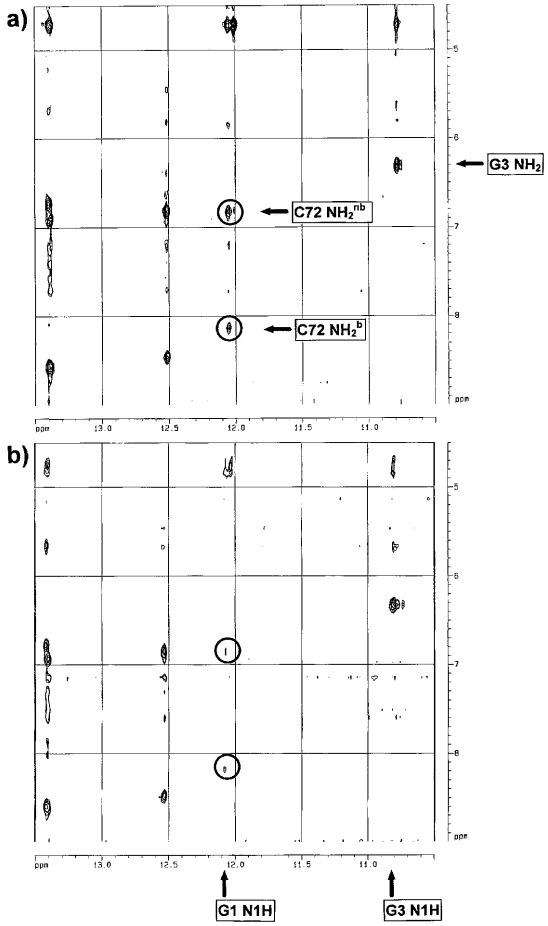

Figure 5. Part of the JR-NOESY spectra (mixing time 300 ms) of the $tRNA^{Ala}$ acceptor stem duplex (concentration ~1.3 mM) at a temperature of 303 K. (**a**) Acceptor helix without protein; (**b**) acceptor helix after addition of 2.3 mg AlaN470 (concentration ~0.09 mM). In both cases the solution contained 100 mM NaCl, 4 mM $MgCl_2$, 10 mM sodium phosphate, pH 6.5. The resonance positions of the G1 and G3 imino protons (horizontal axis) and the C72 and G3 amino protons (vertical axis) are marked. Upper indices nb and b denote the amino protons which are not involved (non-bound) and involved (bound) respectively in Watson–Crick base pairing. Crosspeaks which are distinctly weakened upon addition of AlaN470 are circled.

acceptor helix RNA solution. A small but nevertheless distinct and specific increase in the linewidths is observed, indicating weak binding of the RNA to the protein. The imino signal originating from base pair 2 is broadened by ~5 Hz (from ~12 to ~17 Hz), the linewidths of the imino resonances due to G3–U70 base pair 3 are increased by ~3 Hz (U imino peak; the linewidths of the partially overlapping resonances in the complex of U70 and G1 have been derived from graphical deconvolution of the sum peak) and 9 Hz (G, from ~10 to 19 Hz) respectively, whereas the base pair 1 line is broadened by ~12 Hz (from ~23 to ~35 Hz). No assertions can be made concerning base pairs 4–7. Due to rapid imino proton exchange caused by fraying of terminal base pairs 6 and 7, the corresponding imino resonances of U6 and U66 are not observable at 310 K. Moreover, there is severe overlap of the imino resonances of G4 and G68 at ~13.4 p.p.m. (base pairs 4 and 5, respectively), which prevents a reliable assessment of the effect of enzyme addition in terms of linewidth and chemical shift changes. Nevertheless, slight linewidth increases of both signals are probable and compatible with the observed overall lineshape. Obviously, the degree of broadening is quite specific for the individual imino resonances, pointing to a specific mode of interaction between the acceptor duplex and AlaN470.

This assumption is corroborated by a comparison of the NOESY spectra of the acceptor helix alone (Fig. 5a) and a complex of the same RNA (~1.3 mM) with ~0.09 mM AlaN470 (i.e. a molar RNA:protein ratio of ~14:1). In the NOESY spectrum of the complex (Fig. 5b) the two crosspeaks between the G1 imino resonance and the amino signals of the base paired C72 (at ~8.1 p.p.m. for the proton involved in the Watson–Crick base pairing, at ~6.8 p.p.m. for the non-hydrogen bonded amino proton) are drastically weakened. This can be explained by an increase in the linewidths of both the G1 imino and the C72 amino resonances, which in turn can be attributed to an enhanced exchange rate of these labile protons with the solvent molecule (H_2O) protons.

Interestingly, the crosspeak between G3 imino proton resonance (~10.8 p.p.m.) and the G3 amino proton signal (~6.3 p.p.m.) remains unaffected upon addition of AlaN470 to the RNA solution. The mere observation of this crosspeak seems remarkable, since G amino protons cannot usually be observed in proton NMR spectra due to their comparatively fast exchange rates.

DISCUSSION

Advantage and properties of thermostable AlaRS and AlaN470

Though AlaRS is one of the most interesting and biochemically extensively characterized aaRSs, structural data have not been

available to date. The reason is probably the low stability of this enzyme from mesophilic organisms, which renders difficult both crystallization and production of the homogeneous, highly concentrated solutions that are needed for NMR studies. Numerous examples of stable, well crystallizing proteins from *T.thermophilus* prompted us to clone its gene and overproduce AlaRS from this extreme thermophilic organism. Indeed, the protein proved rather stable and could be purified without major complications. Fortunately, the genetically engineered N-terminal 53 kDa fragment of AlaRS (AlaN470), which is sufficient for specific aminoacylation of cognate $tRNA^{Ala}$, proved as stable as the entire enzyme and could also be purified in greater amounts. This protein is, due to its reduced size, particularly valuable for NMR investigations.

Primary sequence alignment of the *T.thermophilus* AlaRS with other known AlaRSs revealed, as expected, a high degree of similarity around the three characteristic motifs of class II aaRSs. However, a retroviral-like zinc binding motif (C–X_2–C–X_6–H–X_2–H) located between motifs 2 and 3 in the *E.coli* AlaRS (29) was not found at the homologous position of the *T.thermophilus* AlaRS (amino acids 173–186), a property shared with the known eukaryotic AlaRSs. Only one amino acid, namely the second cysteine (Cys176), of this zinc binding motif is present in the *T.thermophilus* AlaRS, while the first cysteine is replaced by a serine. The two histidines completing the motif in the *E.coli* enzyme are not found in the *T.thermophilus* counterpart, whose sequence is rather different in that region. It is, therefore, unlikely that the *T.thermophilus* AlaRS binds a metal atom in this region. For the *T.thermophilus* AlaRS as well as for the eukaryotic AlaRSs a protein-bound metal may not be a requirement for tRNA recognition, in contrast to the corresponding *E.coli* enzyme, as demonstrated by mutagenesis and binding studies (6,7). However, a cysteine (Cys666) that is not involved in zinc binding in the *E.coli* AlaRS (28) is present at the homologous position in the *T.thermophilus* enzyme (amino acid 682). Though located outside the N-terminal fragment (461N), which is sufficient for specific aminoacylation, this residue was shown to contribute to the catalytic efficiency of *E.coli* AlaRS. The whole region (*T.thermophilus* AlaRS amino acids 500–716) where this cysteine is located shows a high degree of sequence similarity with other known AlaRSs, indicating its, albeit unknown, functional importance.

A previous report demonstrated specific DNA binding by the *E.coli* AlaRS (30). It remained unclear which domain was responsible for this property. The *T.thermophilus* phenylalanyl-tRNA synthetase, which belongs to the same subclass (IIc) of aaRSs, also specifically binds DNA fragments located upstream of its own genes (our unpublished results). A sequence comparison of the assumed DNA binding domain 5 of the β-subunit of this enzyme (31) with regions of unknown function of the AlaRS, however, revealed no obvious similarities. The structural basis of this rather unusual and poorly understood common property of both aaRSs has to be established.

A high degree of sequence similarity with other known AlaRSs was also found at amino acids 400–411, which is in the region that probably recognizes the critical G3–U70 base pair of $tRNA^{Ala}$. This region is supposed to form a module that folds back to the catalytic center to ensure proper tRNA recognition by making contacts with its acceptor stem (8). The accumulated conserved aromatic amino acids in this region may interact with the acceptor stem nucleotides.

The AlaRS from *T.thermophilus* was found to be a dimer. Recently it was shown that this also applies to the *E.coli* AlaRS (32). In an earlier publication it appeared as a tetramer (4). Association of *E.coli* AlaRS dimers to form larger species seems to occur only at lower temperatures. For the thermostable *T.thermophilus* AlaRS, however, even at room temperature, which is ~40°C below its optimum, no such association was found. Thus, the thermophile AlaRS resembles the homodimeric structure of most of the other prokaryotic class II aaRSs.

The observed non-aminoacylation by AlaN470 of the C3–A70 acceptor stem duplex *in vitro* is at variance with results obtained with the *E.coli* AlaRS (10), where alanylation of mutant suppressor $tRNA^{Ala}$ containing a C3–A70 mismatch pair was found *in vivo*. The differing behavior might be associated with the strongly reduced binding strength of the acceptor helix as compared with intact $tRNA^{Ala}$. The weakened interaction might prevent induced adaptation (‘fine adjustment’) of the AlaRS contact region to the recognition region of the RNA (mismatch pair), for which tight binding of the RNA could be an essential prerequisite, in particular if the recognition element has been modified with respect to the wild-type sequence.

NMR analyses

In our preliminary NMR investigations presented here the acceptor helix of $tRNA^{Ala}$ from *E.coli*, comprising 7 bp and the single-stranded ACCA 3′-terminus (inset of Fig. 4), was used to study the interaction with AlaN470. Interaction between the tRNA microhelix and AlaN470 is distinctly weaker as compared with the native, original molecule, since both interaction partners have been truncated. For NMR analysis this was advantageous, since the expected reduced association constant gave rise to a fast exchange of the RNA between the free and complexed forms, thus resulting in well-resolved NMR spectra of the RNA. In the case of a tightly bound complex, however, the effective molecular weight of ~60 kDa would have rendered a detailed NMR analysis impracticable, due to the greatly increased linewidths for the RNA moiety and the correspondingly poor spectral resolution.

From the preliminary NMR studies a weak, but specific, interaction between the acceptor stem of $tRNA^{Ala}$ and AlaN470 can be derived. The most pronounced effect in the imino and amino resonance regions of the proton NMR spectra is found for the first base pair (G1–C72), indicating a specific contact of this site with the protein. Thereby, the exchange rate of the corresponding labile imino and amino protons is enhanced in comparison with the non-complexed state. This might be due to a widening or even disruption of the first base pair or specific contact with a side chain of the protein which is able to catalyze proton exchange upon interaction of the acceptor stem with AlaN470 in the major groove. Only slight effects upon the exchangeable proton resonances of the identity element base pair G3–U70 could be observed in the presence of AlaN470. In particular, the G3 amino–G3 imino crosspeak is not influenced by addition of AlaN470 to the acceptor helix solution. The presence of this crosspeak is all the more noteworthy as the G3 amino group is not involved in Watson–Crick base pairing and protrudes into the minor groove. Although it should thus be even more easily accessible to solvent molecules than in the case where the NH_2 group is involved in hydrogen bonding, the opposite is found. In the present case a drastic reduction in amino proton

exchange due to particular protection of this group is detected, maybe by a tightly bound water molecule, as found in a crystal structure analysis of a G–U pair-containing mismatch duplex (33). The lack of major changes in the exchangeable proton resonances cannot be taken as evidence for a lack of any specific interaction between the involved nucleotides and the protein and, in particular, between the G3-NH_2 group and functional groups of the enzyme (9), since here only the exchangeable proton resonances were analyzed. Nevertheless, a change in the hydrogen bonding pattern of the G3 amino group seems improbable in view of the reported results. The linewidth (and, accordingly, linewidth changes) of imino resonances are mainly determined by the exchange rates of the corresponding imino protons with the solvent (H_2O). The exchange rate can be affected by the accessibility to exchange-catalyzing buffer ions (34), by the presence of proper amino acid side chains of the protein or by alterations of the hydrogen bond (base pair) geometry. This can be induced by interaction with the protein and in turn gives rise to altered solvent accessibility. Consequently, from these analyses no assertions about possible changes in the helical geometry, in particular in the vicinity of the G3–U70 identity element, upon interaction with AlaN470 can be deduced. It seems possible that specific interactions between certain functional groups of the G3–U70 base pair and the enzyme occur, which, however, do not cause significant variations in the imino proton exchange rates of the exchangeable protons of the identity element base pair. More detailed structural information can be expected from an analysis of the non-exchangeable proton NMR spectra (in particular NOESY), which is presently underway.

ACKNOWLEDGEMENTS

We thank Prof. Dr M.Sprinzl for encouraging the present work and providing facilities, N.Grillenbeck for skillful technical assistance, Dr H.Faulhammer for peptide sequencing, Drs L.Arnold and G.Ott for RNA syntheses and M.Vogtherr and H.Schübel for support with the preparation of NMR samples.

REFERENCES

1 Eriani,G., Delarue,M., Poch,O., Gangloff,J. and Moras,D. (1990) *Nature*, **347**, 203–206.
2 Cusack,S. (1995) *Nature Struct. Biol.*, **2**, 824–831.
3 Schimmel,P. and Ripmaster,T. (1995) *Trends Biochem. Sci.*, **20**, 333–334.
4 Jasin,M., Regan,L. and Schimmel,P. (1983) *Nature*, **306**, 441–446.
5 Davis,M.W., Buechter,D.D. and Schimmel,P. (1994) *Biochemistry*, **33**, 9904–9911.
6 Miller,W.T. and Schimmel,P. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 2032–2035.
7 Miller,W.T. and Schimmel,P. (1992) *Mol. Microbiol.*, **6**, 1259–1262.
8 Buechter,D.D. and Schimmel,P. (1995) *Biochemistry*, **34**, 6014–6019.
9 Francklyn,C., Musier-Forsyth,K. and Schimmel,P. (1992) *Eur. J. Biochem.*, **206**, 315–321.
10 Gabriel,K., Schneider,J. and McClain,W.H. (1996) *Science*, **271**, 195–197.
11 Musier-Forsyth,K., Usman,N., Scaringe,S., Doudna,J., Green,R. and Schimmel,P. (1991) *Science*, **253**, 784–786.
12 Castenholz,R.W. (1969) *Bacteriol. Rev.*, **33**, 476–504.
13 Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) *Methods Enzymol.*, **185**, 60–89.
14 Miller,J. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
16 Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.
17 Limmer,S., Reiser,C.O.A., Schirmer,N.K., Grillenbeck,N.W. and Sprinzl,M. (1992) *Biochemistry*, **31**, 2970–2977.
18 Lechler,A., Keller,B., Hennecke,H. and Kreutzer,R. (1996) *Protein Expression Purificat.*, **8**, 347–357.
19 Vornlocher,H.-P., Scheible,W.-R., Faulhammer,H.G. and Sprinzl,M. (1997) *Eur. J. Biochem.*, **243**, 66–71.
20 Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Smith,J.A., Seidman,J.G. and Struhl,K. (1987) *Current Protocols in Molecular Biology*. Greene Publishing Associates, New York, NY.
21 Kunkel,T.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 488–492.
22 Zubay,G. (1966) *Procedures in Nucleic Acid Research*. Harper & Row, New York, NY.
23 Fersht,A. (1985) *Enzyme Structure and Mechanism*. W.H.Freeman and Company, New York.
24 Ott,G., Arnold,L., Smrt,J., Sobkowski,M., Limmer,S., Hofmann,H.-P. and Sprinzl,M. (1994) *Nucleosides Nucleotides*, **13**, 1069–1085.
25 Arnold,L., Smrt,J., Zajicek,J., Ott,G., Schiesswohl,M. and Sprinzl,M. (1991) *Collect. Czech. Chem. Commun.*, **56**, 1948–1956.
26 Hore,P.J. (1983) *J. Magn. Res.*, **54**, 539–542.
27 Plateau,P. and Guéron,M. (1982) *J. Am. Chem. Soc.*, **104**, 7310–7311.
28 Wu,M.-X., Filley,S.J., Xiong,J., Lee,J.J. and Hill,K.A.W. (1994) *Biochemistry*, **33**, 12260–12266.
29 Miller,W.T., Hill,K.A.W. and Schimmel,P. (1991) *Biochemistry*, **30**, 6970–6976.
30 Putney,S.D. and Schimmel,P. (1981) *Nature*, **291**, 632–635.
31 Mosyak,L., Reshetnikova,L., Goldgur,Y., Delarue,M. and Safro,M.G. (1995) *Nature Struct. Biol.*, **2**, 537–547.
32 Sood,S.M., Slattery,C.W., Filley,S.J., Wu,M.-X. and Hill,K.A.W. (1996) *Arch. Biochem. Biophys.*, **328**, 295–301.
33 Holbrook,S.R., Cheong,C., Tinoco,I. and Kim,S.-H. (1991) *Nature*, **353**, 579–581.
34 Guéron,M. and Leroy,J.-L. (1995) *Methods Enzymol.*, **261**, 383–413.