Structural Analysis of Multifunctional Peptide Motifs in Human Bifunctional tRNA Synthetase: Identification of RNA-Binding Residues and Functional Implications for Tandem Repeats[†,‡]

Eui-Jun Jeong,[⊥] Geum-Sook Hwang,[⊥,§] Kyung Hee Kim,[§] Min Jung Kim,[§] Sunghoon Kim,*[,§] and Key-Sun Kim*[,⊥]

Structural Biology Center, Korea Institute of Science and Technology (KIST), Cheongryang Box 131, Seoul, 130-650, and National Creative Research Initiatives Center for ARS Network, Sung Kyun Kwan University, Suwon, Kyunggido, 440-746, Korea

Received June 19, 2000; Revised Manuscript Received October 10, 2000

ABSTRACT: Human bifunctional glutamyl-prolyl-tRNA synthetase (EPRS) contains three tandem repeats linking the two catalytic domains. These repeated motifs have been shown to be involved in protein−protein and protein−nucleic acid interactions. The single copy of the homologous motifs has also been found in several different aminoacyl-tRNA synthetases. The solution structure of repeat 1 (EPRS-R1) and the secondary structure of the whole appended domain containing three repeated motifs in EPRS (EPRS-R123) was determined by nuclear magnetic resonance (NMR) spectroscopy. EPRS-R1 consists of two helices (residues 679−699 and 702−721) arranged in a helix−turn−helix, which is similar to other RNA binding proteins and the j-domain of DnaJ, and EPRS-R123 is composed of three helix−turn−helix motifs linked by an unstructured loop. When tRNA is bound to the appended domain, chemical shifts of several residues in each repeat are perturbed. However, the perturbed residues in each repeat are not the same although they are in the same binding surface, suggesting that each repeat in the appended domain is dynamically arranged to maximize contacts with tRNA. The affinity of tRNA to the three-repeated motif was much higher than to the single motif. These results indicate that each of the repeated motifs has a weak intrinsic affinity for tRNA, but the repetition of the motifs may be required to enhance binding affinity. Thus, the results of this work gave information on the RNA-binding mode of the multifunctional peptide motif attached to different ARSs and the functional reason for the repetition of this motif.

Aminoacyl-tRNA synthetases (ARSs)[1] catalyze the attachment of specific amino acids to their cognate tRNAs. Although their catalytic activity is conserved in all living organisms, these enzymes are diversified in the structures and functions that are not necessarily essential for the catalysis. Structurally, mammalian ARSs are usually larger than their counterparts in prokaryotes by having noncatalytic extensions. Also several mammalian ARSs are associated to form a macromolecular complex (*1*−*4*). The existence of these structural features in mammalian ARSs implies novel cellular functions that may be specifically required in multicellular organisms. This possibility is supported by the recent findings that human tyrosyl-tRNA synthetase is secreted and cleaved to form two distinct pro-apoptotic cytokines (*5*) and that arginyl-tRNA synthetase is associated with the precursor of a pro-apoptotic cytokine (*6*).

Although it has been known for more than two decades that several eukaryotic ARSs form a multiprotein complex, the physiological significance of this complex or the nature of protein−protein interaction in this complex has not been well understood. This complex consists of at least nine synthetases (IRS, LRS, MRS, QRS, RRS, KRS, DRS, and a bifunctional EPRS) and three auxiliary proteins, p18, p38, and p43 (*7*−*9*). The assembly of this complex appears to be mediated by cooperative protein−protein interactions via different structural domains. The noncatalytic extensions of the complex-forming ARSs are involved in protein−protein interactions (*4*), while p38 functions as a scaffold for the assembly of ARSs (*9*). In addition, even the catalytic region of human QRS is required for the assembly of the multi-ARS complex (*54*). Although the protein−protein interactions between the peptide extensions were shown among the complex-forming ARSs, other ARSs not found in the multi-ARS complex also contain similar extensions. This suggests that these peptide extensions may also play additional novel functions

[†] This work was supported in part by Critical Technology 21 (K.-S.K.) and by a grant from National Creative Research Initiatives (S.K.) of the Ministry of Science and Technology of Korea.

[‡] The atomic coordinates (code 1fyj) have been deposited in the Protein Data Bank at the Research Collaboratory for Structural Bioinformatics.

* To whom correspondence should be sent. Key-Sun Kim, Telephone: (+82-2) 958-5934. Fax: (+82-2) 958-5939. E-mail: keysun@kist.re.kr; Sunghoon Kim, Telephone: (+82-31) 290-5680. Fax: (+82-31) 290-5682. E-mail: sunghoon@yurim.skku.ac.kr.

[⊥] Korea Institute of Science and Technology.

[§] Sung Kyun Kwan University.

[1] Abbreviations: ARSs, aminoacyl-tRNA synthetases; XRS, aminoacyl-tRNA synthetases for amino acid X; NOESY, nuclear Overhauser effect enhancement spectroscopy; TROSY, transverse-relaxation-optimized spectroscopy; HSQC, heteronuclear single-quantum coherence; TOCSY, total correlation spectroscopy.

10.1021/bi001393h CCC: $19.00
Published on Web 11/22/2000

```
EPRS-Hum R1   677-DSLVLYNRVAV QGDVVRELKAKKAPKEDVDAAVKQL LSLKAEYKEKTGQEYKPGNPP AEIGQNISSNSSASIL
EPRS-Hum R2   750-ESKSLYDEVAA QGEVVRKLKAEKSPKAKINEAVECL LSLKAQYKEKTGKEYIPGQPP LSQSSDSSPTRNSEPAGLETP
EPRS-Hum R3   828-EAKVLFDKVAS QGEVVRKLKTEKAPKDQVDIAVQEL LQLKAQYKSLIGVEYKPVSAT
WRS-Hum        12-    LFNSIAT QGELVRSLKAGNASKDEIDSAVKML VSLKMSYKAAAGEDYK
HRS-Hum         3-ERAALEELVKL QGERVRGLKQQKASAELIEEEVAKL LKLKAQ
GRS-Hum        17-        VRQ QGDLVRKLKEDKAPQVDVDKAVAEL
MRS-Hum       841-QIQALMDEVTK QGNIVRELKAQKADKNEVAAEVAKL LDLKKQLAVAEGKPPEAPKGK
```

FIGURE 1: Amino acid sequence of the motifs present in different human ARSs. Human EPRS contains three tandem repeats of this motif (R1−3). The sequence corresponding to EPRS-R1 ($D^{677}-P^{733}$) is underlined.

The motifs in EPRS are unique in a few points. First, they are present as tandem repeats (*10*), while others mostly exist as a single copy. Second, they are located between the two catalytic domains while others are attached to the N- or C-termini of the core domain. Third, the homologous motifs are widely present throughout different ARSs of various species, implicating their functional significance (*11*). Nonetheless, the roles of these motifs are not completely clear, although they are expected to be multifunctional, being involved in protein−protein and protein−nucleic acid interactions (*11*, *12*). The motifs homologous to the EPRS repeats have been found as a single copy in the N-terminal extensions of glycyl- (*13*), histidyl- (*14*) and tryptophanyl-tRNA synthetases (*15*) and in the C-terminal region of methionyl-tRNA synthetase (*4*, *16*). Interestingly, these motifs were identified as an antigenic epitope for autoantibodies detected in myositis patients (*16*). Human histidyl-tRNA synthetase lacking this peptide region lost its antigenicity and catalytic activity, suggesting its importance in these two activities (*17*). Recently, the tertiary structure of this motif from the hamster EPRS has been reported (*18*). However, the binding properties of this motif to tRNA or protein have not been elucidated. In this work, we have determined the solution structure of the 57 amino acid motif present in the human EPRS as tandem repeats and mapped the binding site of tRNA in repeated motif 1 (EPRS-R1) and the whole appended domain (EPRS-R123).

EXPERIMENTAL PROCEDURES

Preparation of the Repeated Motif of EPRS. A repeated motif containing residues $Asp^{677}-Pro^{733}$ (EPRS-R1, Figure 1) was subcloned into an expression vector pET28a (Novagen) by PCR amplification from the human EPRS gene with primers containing restriction enzyme sites, *Eco*RI and *Xho*I. The EPRS-R1 containing poly-histidine tag at the N-terminus was expressed in *Escherichia coli* BL21 (DE3). The fragment, EPRS-R123, containing all three repeated motifs ($Asp^{677}-Thr^{884}$) was subcloned into an expression vector pET15b. The transformed cells were cultured at 37 °C and induced with 0.4 mM IPTG (isopropyl-1-thio-β-D-galactopyranoside) at the logarithmic phase of growth for 3 h. The harvested cell paste was disrupted by an ultrasonicator (Branson) in a lysis buffer (50 mM sodium phosphate, pH 8.0, and 300 mM NaCl). The protein was purified by an affinity chromatography with nickel-nitrilotriacetic acid-agarose column (Qiagen). The polyhistidine tag was removed by subtilisin or thrombin digestion, and finally purified by reversed-phase chromatography using a C_8 Vydac column. Purified protein was lyophilized and stored at −20 °C. A uniformly ^{15}N-labeled protein was prepared from cells grown in M9 minimal medium containing 1.0 g of $^{15}NH_4Cl$/L and 2.0 g of glucose/L, and glucose was substituted by ^{13}C-glucose for the simultaneous isotope labeling of ^{13}C and ^{15}N. NMR samples were prepared by dissolving about 2 mM protein in 0.5 mL of either 90% H_2O/10% ^2H_2O or 99.9% ^2H_2O, and the pH was adjusted to 5.0 ± 0.05 (glass electrode, uncorrected) with concentrated NaO^2H. Purified protein was verified by N-terminal amino acid sequence at Korea Basic Science Institute (KBSI) in Seoul. The expressed protein of EPRS-R1 was cleaved for 10 min at 0 °C with subtilisin instead of thrombin to remove other flanking residues in addition to the histidine-tag. The N-terminal sequence of cleaved protein showed that the major portion of protein has a two-residue cloning artifact (Glu−Phe), and the minor portion has a six-residue cloning artifact (Gly−Arg−Gly−Ser−Glu−Phe). The minor fraction of N-terminal extended protein does not confuse the cross-peak assignments since the chemical shifts of both forms were all the same except a few N-terminal residues. Considering that subtilisin is a nonspecific protease, protease digestion indicates that EPRS-R1 retains a tightly packed conformation. EPRS-R123 has the N-terminal extension of Gly−Ser−His−Met.

NMR Spectroscopy. NMR experiments were carried out using a Varian UNITYPlus 600-MHz spectrometer (at the Advanced Analysis Center of KIST) at 30 °C. Two-dimensional ^1H-DQF-COSY (*19*), ^1H-TOCSY (*20*), and ^1H-NOESY (*21*) were acquired with a plain sample in water and three-dimensional ^{15}N-edited TOCSY, ^{15}N-edited NOESY (*22*), and HNHA (*23*), as well as 2D-^1H-^{15}N HSQC were acquired with a ^{15}N-labeled sample. As for the EPRS-R123, 3D-HNCACB (*24*), CBCA(CO)NH (*24*), HNCO (*25*), HNCA (*25*) and ^{13}C-, ^{15}N-edited NOESY (*26*) were acquired additionally. Data were processed and analyzed using NMRPipe (*27*) and PIPP (obtained from Dr. Garrett, NIH), respectively.

Assignments and Distance Restraints. Starting with identifications of ^{15}N and HN chemical shifts in ^1H-^{15}N HSQC, spin systems were partly identified in ^{15}N-edited TOCSY, and the sequential assignments of each amino acid were made using ^{15}N-edited NOESY. Stereospecific assignments of Hβ protons were based on the intensity of HN-Hβ cross-peaks in ^{15}N-edited TOCSY and ^{15}N-edited NOESY spectra (*28*). The methyl groups of Val and Leu were stereospecifically assigned based on the intensity of HN-Hγ, Hα-Hγ cross-peaks, and NOE intensity of the sterospecifically assigned Hβ protons to δ-methyl protons of Leu. NOE distance restraints were derived from 3D ^{15}N-edited NOESY in H_2O and 2D ^1H NOESY spectra in ^2H_2O, all with a mixing time of 150 ms. All NOE cross-peaks were assigned using a program PIPP. The NOE intensity was converted into three groups of classes (1.8−2.7, 1.8−3.5, and 1.8−5.0) and pseudo-atom corrections were made appropriately (*29*). The $^3J_{HN\alpha}$ scalar-coupling constant of the α-proton to the amide proton was obtained from the HNHA experiments. The

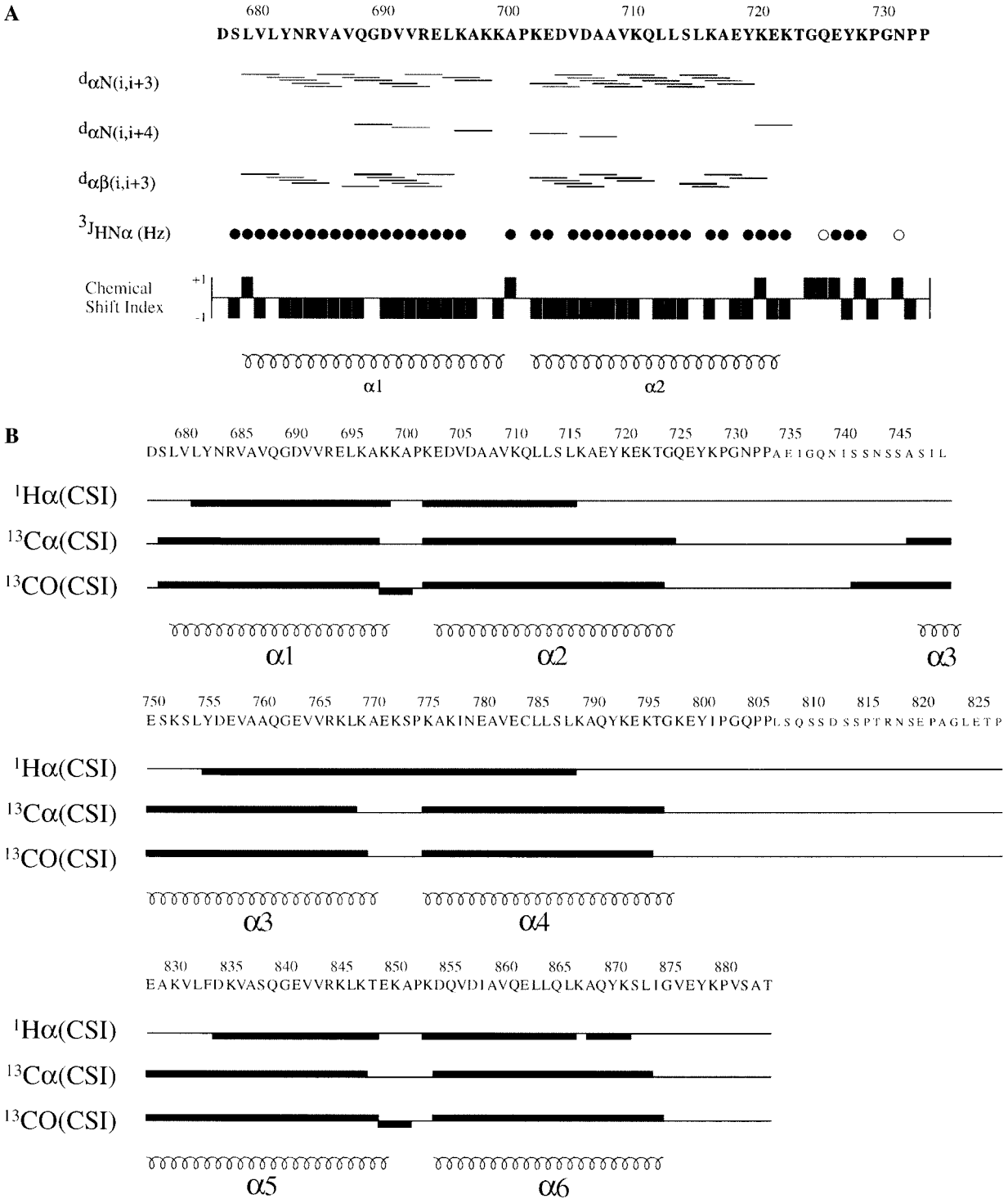

FIGURE 2: Summary of NOE, $^3J_{HNH\alpha}$ coupling constants, and chemical shift index of EPRS-R1 calculated from $^1H\alpha$ chemical shifts (A) and chemical shift indices of EPRS-R123 calculated from $^1H\alpha$, $^{13}C\alpha$, and ^{13}CO chemical shifts (B) according to Wishart et al. (*32*). The $^3J_{HNH\alpha}$ coupling constants either lower than 5.5 Hz (●) or higher than 8 Hz (○) are indicated.

backbone torsion angle ϕ was restrained to −85 to −25 for $^3J_{HNH\alpha}$ <5.5 Hz, −70 to −170 for $^3J_{HNH\alpha}$ 8−9 Hz. The ι torsion angles of helix region were restrained to −70 to −10. The side chain torsion angles χ_1 were restrained based on cross-peak intensities deduced from ^{15}N-edited TOCSY and ^{15}N-edited NOESY spectra (*28*). Additional backbone H-bond restraints were given, where secondary structures were indicated based on NOE connectivity. For each hydrogen bond, two restraints (r_{NH-O}, 1.7−2.3; r_{N-O}, 2.5−3.3) were used. Additionally, $^3J_{HNH\alpha}$-coupling constants (*30*) were directly included in the simulated annealing protocol during refinement. The chemical shifts assignments of EPRS-R123 were made using CBCACONH and HNCACB in comparison with the assignments of EPRS-R1. Assignments were further confirmed by ^{15}N-edited NOESY, ^{13}C-, ^{15}N-edited NOESY, and specific labeling of lysines and leucines of protein. Carbonyl carbons were assigned using HNCA and HNCO experiments.

Structure Calculation. The structures were calculated using a distance geometry and simulated annealing protocol using

an X-PLOR program (*31*). The secondary structures were initially identified with the chemical shift index (*32*), $^3J_{HNH\alpha}$-coupling constants, and medium range NOE connectivity (*33*). In the initial structure calculations from an extended chain, NOE restraints and hydrogen bonds in the helix region were included. When the initial structures were obtained, $^3J_{HNH\alpha}$-coupling constants and dihedral angles were included. During the calculations, ambiguous assignments were clarified and the refinements of structures were achieved with the inclusion of database potential (*34*, *35*).

NMR Analysis of the tRNA Binding. The binding of yeast tRNA to EPRS motif was monitored by measuring chemical shift perturbation upon adding tRNA to EPRS-R1 or EPRS-R123 polypeptide at the same molar ratio in the presence of 150 mM NaCl. The ^{1}H-^{15}N HQSC spectra implemented with TROSY (*36*) of EPRS-R1 or EPRS-R123 were acquired with or without tRNA at pH 6.0. The acquired spectra were processed with the same processing parameters and were overlaid on each other to identify the residues with chemical shift perturbation. The sample was adjusted to pH 6.0 because the yeast tRNA was not soluble at a lower pH than 6.0. Only backbone amide nitrogens of leucines and lysines of EPRS-R123 are specifically labeled.

Surface Plasmon Resonance Analysis. The interaction of the EPRS repeats with tRNA was analyzed by surface plasmon resonance analysis (BIAcore 2000, Amersham Pharmacia Biotech). The DNAs encoding EPRS-R123 and EPRS-R1 were cloned into *EcoR*1 and *Sal*1 sites of pET28a and expressed in *E. coli* BL21 (DE3). The peptides were purified using nickel affinity chromatography following the manufacturer's protocol (Novagen). Each of the purified peptides was immobilized to a CM5 sensor chip using the amine-couple kit, and a continuous flow of running buffer (10 mM Hepes, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% v/v surfactant P20) was maintained at 20 μL/min for all of the experiments. The sensor chip surface was activated using 35 μL of 0.05 M *N*-hydroxysuccinimide and 0.2 M *N*-ethyl-*N*′-(dimethylaminopropyl) carboimide. Proteins diluted in the running buffer were applied to the sensor surface at the flow rate of 5 μL/min to create a covalent link through their primary amine group. The remaining active sites on the surface were blocked and deactivated with 35 μL of 1.0 M ethanolamine HCl, pH 8.5. EPRS-R123 (0.75 mg/mL), EPRS-R1 (0.97 mg/mL), and BSA (1 mg/mL) were immobilized on the surface to get about 4000RU, 1000RU, and 5000RU, respectively. Bovine tRNA as an analyte was diluted to various concentrations with the running buffer and was then injected over the sensor chip surface for 2 min to record the association. The running buffer was followed without a delay after the analyte solution to measure the dissociation. The chip surfaces were regenerated at the end of each detection by injecting 10 μL of 25 mM NaOH and 500 mM NaCl. The sensograms were corrected by subtracting the responses of the BSA surface from the responses of peptide-immobilized sensor chip surfaces and were analyzed using BIAevaluation 3.0. All experiments were repeated three times to have reproducible results.

RESULTS

Solution Structure of EPRS-R1. Human EPRS contains three tandem-repeated peptides of 57 amino acids. EPRS-R1 shows the archetypical helix−turn−helix structural motif as shown in Figure 2, panel A. The helical NOEs are observed in residues from 679 to 699 and from 702 to 721. The $^3J_{HN\alpha}$ coupling constants and chemical shift index are also consistent with NOE connectivity. The three-dimensional structure of EPRS-R1 (residues Asp^{677}−Pro^{733}) is defined by 814 experimentally derived NOE distance restraints, 91 dihedral angle restraints, 29 hydrogen bonds, and 49 $^3J_{HNH\alpha}$ coupling constants. The root-mean-square deviation for backbone and all heavy atoms of 20 minimized structures around mean atomic displacements of the whole molecule (residues Asp^{677}−Pro^{733}) was 0.96 and 1.39 Å, respectively, and that of well-structured regions (residues 679−721) are 0.48 and 1.0 Å, respectively. More than 85% of residues are in the most favorable region in Ramachandran plot (Table 1) with no violation of NOE restraints greater than 0.5 Å and dihedral angles greater than 5°. Two helices (679−699, 702−721) interact each other with hydrophobic residues (Val^{692}, Leu^{695}, Ala^{670}, Val^{705}, and Ala^{708}) proximal to the turn involved in helix−helix interactions. However, helices distal to the turn do not have intense interactions between them weakening the helix−helix formation. Instead, the C-terminal loop folds back to interact with the aromatic residues (Tyr^{682} and Tyr^{719}) in the helices (Figure 3, panel a). The hydrophobic interactions between Tyr^{727} in the loop and Tyr^{682} and Tyr^{719} in the helices are crucial for the C-terminal loop formation. Some hydrophobic residues are involved in helix−helix interactions, but others are exposed on the surface, implying that they may play a role in protein−

Table 1: Summary of Structural Restraints Derived from NOEs, Coupling Constants

NOE distance restraints	
intraresidual	394
sequential	175
medium-range($1 < \|i - j\| = 4$)	125
long-range ($\|i - j\| = 5$)	120
hydrogen bonds	2*29[a]
angular restraints	
dihedral angle ϕ	41
dihedral angle ψ	33
dihedral angle χ^1	17
coupling constants	49
Ramachandran plot[b] (residues 3−45)	
most favorable region	92.7
additionally allowed region	7.3
generously allowed region	0.0
disallowed region	0.0
atomic rmsd values (Å)	residues 679−721 (677−733)[c]
backbone atoms	0.49 (0.96)[c]
heavy atoms	1.10 (1.39)[c]
X-PLOR energy terms (kcal/mol)	
E_{tot}	−260.55 ± 44.25
E_{bond}	13.79 ± 0.77
E_{angle}	116.32 ± 4.75
$E_{improper}$	17.91 ± 1.25
E_{vdw}	−467.90 ± 41.78
E_{cdih}	0.03 ± 0.10
E_{NOE}	59.31 ± 7.67

[a] For each hydrogen bond, two distance restraints are used: r_{HN-O} 1.7−2.3, r_{N-O} 2.5−3.3. Coupling constants are used for restraints based on Clore and co-workers (*34*). Coupling constants are values obtained from HNHA experiments. [b] The program PROCHECK_nmr (*53*) was used to analyze the quality of the structure. The values for X-PLOR energy terms were obtained with force constraints of 4 kcal mol^{-1} Å^{-4} (E_{vdw}), 50 kcal mol^{-1} Å^{-2} (E_{NOE}), and 200 kcal/mol (E_{cdih}). E_{vdw} energy is L−J energy of XPLOR (*31*) energy terms. [c] Values for residues 677−733 are in parentheses.

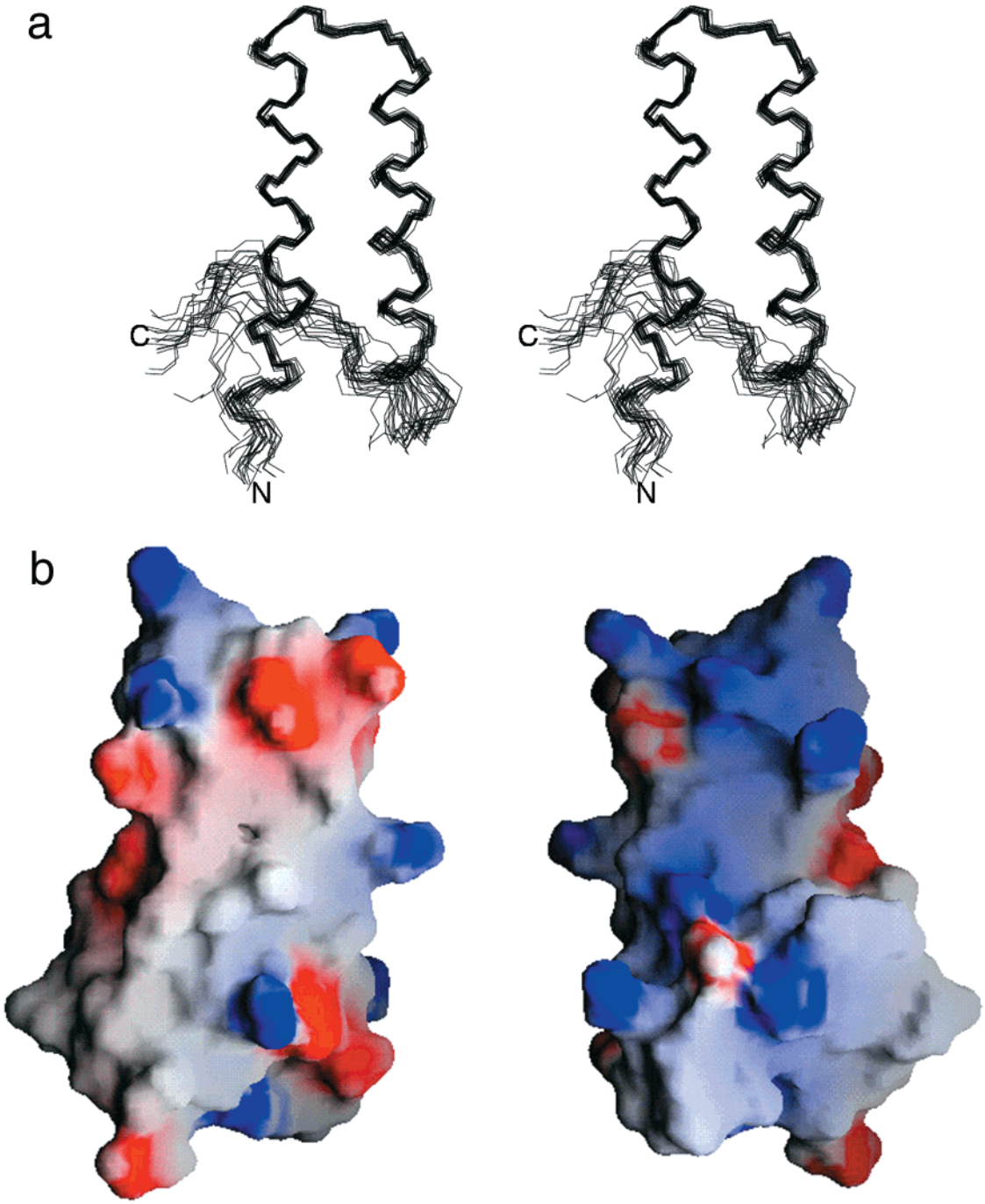

FIGURE 3: Structure of EPRS-R1. (a) Stereoview of the backbone atoms (N, Cα, C′) of 20 structures of EPRS-R1. The root-mean-square deviation about the mean coordinate position for residues 3−45 is 0.48 Å for the backbone atoms and 1.0 Å for all heavy atoms. No distance restraints are violated more than 0.5 Å in any structures, and no torsion angle restraints are violated more than 5°. Other structural statistics are summarized in Table 1. (b) Surface electrostatic potential is color-coded. The negative potential surface is in red ($<-8k_BT$), and the positive potential surface is in blue ($>8k_BT$). The orientation of the figure in the left is the same as in Figure 3, panel A, and the opposite side of electrostatic potential surface is in the right. Electrostatic potential surface was generated by the program GRASP (*49*).

protein or protein−nucleic acid interactions. The relatively well-buried residues with less than 10% of the average solvent-accessible surface area are Val^{685}, Ala^{686}, Val^{692}, Val^{705}, and Ala^{708}, and residues with the average solvent-accessible surface area of 10−20% are Tyr^{682}, Gln^{688}, Gly^{689}, Leu^{695}, Ala^{700}, Leu^{712}, Leu^{715}, and Tyr^{719}. The two surfaces of the helix−turn−helix motif are very different: one side that has interactions with the C-terminus is positively charged and the other side is relatively hydrophobic (Figure 3, panel b).

The previously reported structure of EPRS repeat 2 [PDB code 1R1b (*18*)] from hamster is similar to our structure with the backbone root-mean-square deviation of 1.3 Å in the well-structured region (residues 689−719). However, the structures of the N- and C-terminal regions are different, which may indicate that functional interactions of the repeated motif occur in the region proximal to the turn.

EPRS-R123 contains 208 amino acids and consists of three repeats of the 57 amino acid motif and linkers between them (Figure 1). The secondary structure estimated by chemical

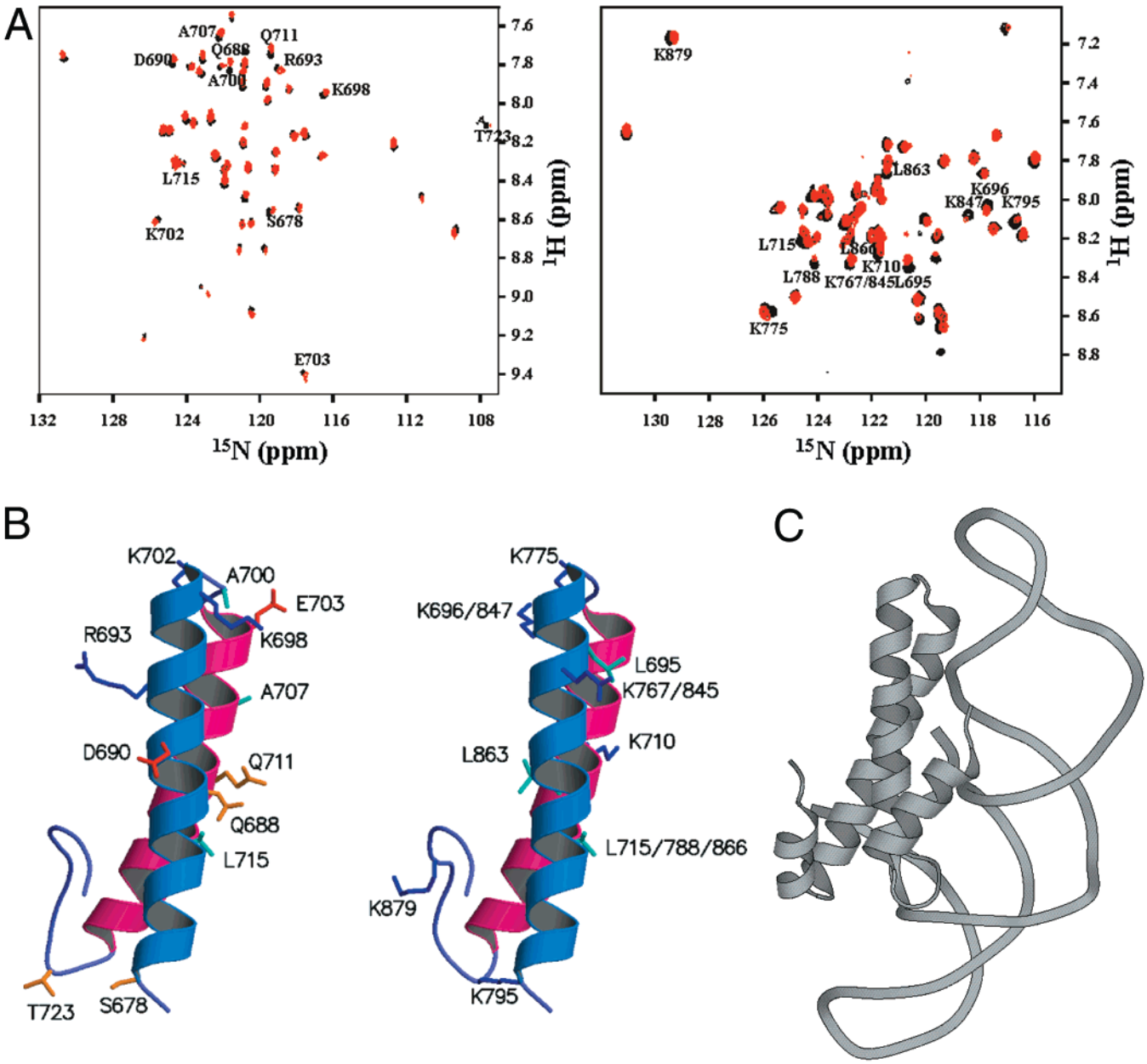

FIGURE 4: Binding of tRNA to the EPRS motif. (A) The binding site of tRNA was mapped by chemical shift perturbation. The cross-peaks of EPRS alone in ^{15}N-^{1}H HSQC spectra are shown in black, and perturbed cross-peaks with tRNA are shown in red. Residues with the weighted chemical shift perturbation ($\delta^{1}H + \delta^{15}N/5$) (*50*) higher than 0.04 ppm are labeled in EPRS-R1 (left) and 0.03 ppm in EPRS-R123 (right). (B) The residues perturbed by tRNA binding to EPRS-R1 (left) or EPRS-R123 (right) are shown in EPRS-R1 structure assuming that the structure of each repeat is similar to that of EPRS-R1. (C) The complex structure of ribosomal protein S15 and ribosomal RNA (*44*). The complex structure shows that the EPRS repeat and ribosomal protein S15 bind to RNA in similar pattern. Ribbon drawing was generated by programs MOLSCRIPT (*51*) and RASTER3D (*52*).

shift index (*32*) shows that each repeat has a helix−turn−helix structural motif similar to the structure of EPRS-R1 and connected by the unstructured loop between them (Figure 2, panel B).

Interactions of EPRS Motif with tRNA. Previously, we showed that the EPRS motif bound to nucleic acids as well as to the IRS motif (*12*). To investigate the binding pattern of the EPRS motif to tRNA, we titrated EPRS-R1 or EPRS-R123 with yeast tRNA. Residues showing chemical shift perturbation induced by tRNA binding are shown in Figure 4, panel A. In EPRS-R1, the chemical shifts of many residues are perturbed. However, most of the perturbed residues exist proximal to the turn on one side of surface that does not have the C-terminal interaction (Figure 4, panel B). Among the specifically labeled lysines and leucines of EPRS-R123, lysines of 696, 710, 767/845, 775, 795, 847, and 879, and leucines of 695, 715, 788, 863, and 866 are perturbed by binding to tRNA (Figure 4, panel A). Lysines and leucines exist throughout the sequence, but not all residues were perturbed by binding to tRNA, and the perturbed residues are mainly on one side of the surface assuming that the structure of each repeat is similar to EPRS-R1 (Figure 4, panel B). In Figure 4, panel B, the perturbed leucines and lysines are shown assuming that the structure of each repeat of EPRS-R123 is the same as the EPRS-R1 structure. Lysines in repeat 2 (K767) and 3 (K845) are substituted for glutamic acid at the equivalent positions in EPRS-R1 (Figure 1). The perturbed surface is the same both in EPRS-R1 and in EPRS-

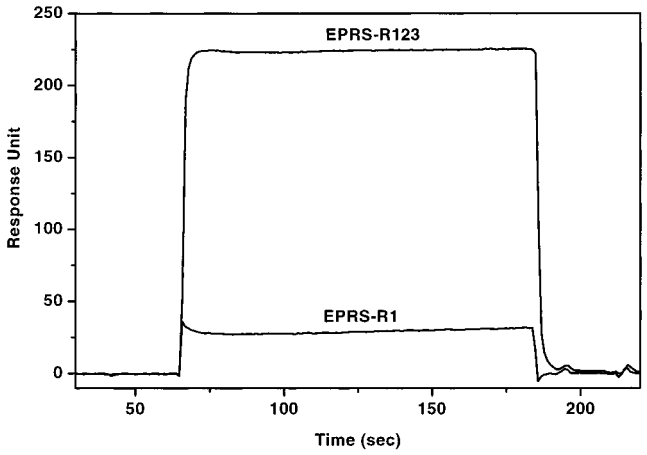

FIGURE 5: Sensogram of real time binding of tRNA to EPRS. The EPRS-R123 and EPRS-R1 were immobilized on the surface to get about 4000 and 1000, respectively, and a flow of tRNA in running buffer at concentration of 50 μM was maintained over the protein-coupled chip.

R123. However, the perturbed residues are not at the same positions as aligned in Figure 1 indicating that a certain arrangement of each motif is needed to maximize the contact with tRNA.

The real time binding analysis shows that the binding of tRNA to ERPS-R1 was too weak to obtain a reliable affinity constant (Figure 5). However, the binding of tRNA to EPRS-R123 is substantially increased.

DISCUSSION

Structure of the EPRS Motif. Most of mammalian ARSs contain peptide appendices in their noncatalytic regions, although their location and length are idiosyncratic. The majority of them are predicted to form helical structures (*3*). Here, we have determined the solution structure of the 57-amino acid motif of EPRS. Although this motif is present as tandem repeats in human EPRS (*10*), *Drosophila* EPRS (*37*), and nematode ERS (*38*), the single copy of the homologous motifs has also been found in HRS (*14*), GRS (*13*), WRS (*15*), and MRS (*4*).

The EPRS motif forms a helix−turn−helix structure, and its folding is highly reversible (*12*). When the structure of EPRS-R1 was compared to known structures with the program DALI (*39*), two RNA-binding proteins were detected. One is ROP protein (*40*) that is encoded by ColE1 plasmid of *E. coli* and is involved in the regulation of a plasmid copy number. The other is a ribosomal protein S15 (*41*) from *Bacillus stearothermophilus* that binds to the central domain of 16S rRNA and also regulates its own synthesis by binding to its own mRNA. The structural similarity of EPRS-R1 to these two RNA binding proteins further supports that EPRS binds to nucleic acids as previously verified (*12*). In addition, the helix−turn−helix structure in the j-domain of DnaJ from *E. coli* is similar to the EPRS motif structure, implying that this motif is involved in protein−protein interaction similar to DnaJ (*42*).

In the EPRS-R1 structure, the tRNA binding site mapped by NMR is located mostly on one side of the surface that is relatively hydrophobic as compared to the other side. The other side is more positively charged and has interactions with the C-terminus. Interactions of the helix−turn−helix with the C-terminus appear to prevent the binding of either protein or nucleic acid on the positively charged side of protein, probably by steric hindrance (Figure 3).

The complex structure of the ribosomal protein S15 from *Thermus thermophilus* with 16S rRNA showed that rRNA binds to S15 protein on one side of surface (*43*, *44*). The helix α4 of S15 protein interacts with helices α2 and α3 in a pattern similar to the C-terminal loop of EPRS-R1 (Figure 4, panel C). This suggests that each repeat of EPRS-R123 interacts with tRNA on one side of protein similar to S15 protein, but they are dynamically arranged to maximize binding contacts with tRNA. This notion was supported by the different chemical shift perturbations in each repeat by tRNA. The fact that the EPRS motif binds to calf thymus DNA or tRNA but not to single-stranded poly(dA)$_{30}$ (*12*) implies that the EPRS motif is a nonspecific nucleic acid-binding motif recognizing the structure of nucleic acids. Since the bound rRNA to S15 protein forms double helix and tRNA also has double helix character, we think EPRS motif recognizes double helix. However, weak binding by general electrostatic interactions with any type of nucleic acids is still possible.

The tandem repeats of the EPRS motif enhanced the binding affinity to tRNA dramatically (Figure 5). The secondary structure of each repeat in the EPRS motif is similar (Figure 2, panel B), and the chemical shifts of EPRS-R1 alone and in tandem repeats are almost overlapped except for a few residues in the N-terminus (data not shown) implies that each repeat is independent one another. Considering that each repeat of EPRS is independent, the repetition of the appended motifs appears to be required to play multiple roles in protein−protein or protein−nucleic acid interactions with higher flexibility and efficiency. As indicated in real time binding of EPRS-R1 and EPRS-R123, each repeat has a weak binding affinity for tRNA but each repeat is dynamically arranged to maximize contacts with tRNA resulting in a high affinity binding.

Biological Implications for the EPRS Motif. Various biological functions have been implicated for the noncatalytic motifs attached to the catalytic domains of tRNA synthetases including the repeated motifs of EPRS. One is that they are involved in protein−protein interactions for the multi-tRNA synthetase complex (*4*, *6*, *11*, *45*, *46*) and another is that they are used as a tRNA-binding motif for delivery and efficiency (*12*, *47*, *48*). Our results showed that the EPRS motif binds to tRNA with one side of the protein and the structure of the EPRS motif was similar to both RNA binding proteins and j-domain of DnaJ, which implies that this structural motif is involved in both protein−nucleic acid and protein−protein interaction. Interactions of the EPRS motif with nucleic acids and proteins suggest that the interaction of the EPRS motif with tRNA may interfere with its association with other tRNA synthetases, increasing the accessibility to the active sites of EPRS that may be buried within the multi-synthetase complex.

In summary, we have determined the solution structure of the multifunctional motif present in human EPRS as well as in several other tRNA synthetases. The RNA binding mode of this helix−turn−helix motif is also found in other homologous RNA binding motifs, and thus it may be commonly used in protein−RNA interactions, and also a similar structure was found in j-domain of DnaJ. We also

found that the repetition of this helix−turn−helix motif is to enhance functional flexibility and binding affinity in the molecular interactions.

SUPPORTING INFORMATION AVAILABLE

The strip plot of the CBCA(CO)NH and HNCACB spectra of EPRS-R123 for the residues L131 to D136. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Mirande, M. (1991) *Prog. Nucleic Acid Res. Mol. Biol. 40*, 95−142.
2. Kisselev, L. L., and Wolfson, A. D. (1994) *Prog. Nucleic Acid Res. Mol. Biol. 48*, 83−142.
3. Yang, D. C. (1996) *Curr. Top. Cell Regul. 34*, 101−36.
4. Rho, S. B., Kim, M. J., Lee, J. S., Seol, W., Motegi, H., Kim, S., and Shiba, K. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 4488−93.
5. Wakasugi, K., and Schimmel, P. (1999) *Science 284*, 147−51.
6. Park, S. G., Jung, K. H., Lee, J. S., Jo, Y. J., Motegi, H., Kim, S., and Shiba, K. (1999) *J. Biol. Chem. 274*, 16673−6.
7. Quevillon, S., and Mirande, M. (1996) *FEBS Lett. 395*, 63−7.
8. Quevillon, S., Agou, F., Robinson, J. C., and Mirande, M. (1997) *J. Biol. Chem. 272*, 32573−9.
9. Quevillon, S., Robinson, J. C., Berthonneau, E., Siatecka, M., and Mirande, M. (1999) *J. Mol. Biol. 285*, 183−95.
10. Fett, R., and Knippers, R. (1991) *J. Biol. Chem. 266*, 1448−55.
11. Rho, S. B., Lee, K. H., Kim, J. W., Shiba, K., Jo, Y. J., and Kim, S. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10128−33.
12. Rho, S. B., Lee, J. S., Jeong, E.-J., Kim, K.-S., Kim, Y. G., and Kim, S. (1998) *J. Biol. Chem. 273*, 11267−73.
13. Ge, Q., Trieu, E. P., and Targoff, I. N. (1994) *J. Biol. Chem. 269*, 28790−7.
14. Tsui, F. W., and Siminovitch, L. (1987) *Nucleic Acids Res. 15*, 3349−67.
15. Frolova, L., Sudomoina, M. A., Grigorieva, A., Zinovieva, O. L., and Kisselev, L. L. (1991) *Gene 109*, 291−6.
16. Hirakata, M., Suwa, A., Takeda, Y., Matsuoka, Y., Irimajiri, S., Targoff, I. N., Hardin, J. A., and Craft, J. (1996) *Arthritis Rheum. 39*, 146−51.
17. Raben, N., Nichols, R., Dohlman, J., McPhie, P., Sridhar, V., Hyde, C., Leff, R., and Plotz, P. (1994) *J. Biol. Chem. 269*, 24277−83.
18. Cahuzac, B., Berthonneau, E., Birlirakis, N., Guittet, E., and Mirande, M. (2000) *EMBO J. 19*, 445−52.
19. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−85.
20. Davis, D. G., and Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820−2821.
21. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
22. Zhang, H., Zhao, D., Revington, M., Lee, W., Jia, X., Arrowsmith, C., and Jardetzky, O. (1994) *J. Mol. Biol. 238*, 592−614.
23. Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) *J. Biomol. NMR 4*, 871−8.
24. Muhandiram, D. R., Kay, L. E. (1994) *J. Magn. Reson. Series B. 103*, 203-216.
25. Ikura, M., Kay, L. E., and Bax, A. (1990) *Biochemistry 29*, 4659−67.
26. Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1994) *J. Magn. Reson. Series B 103*, 197−201.
27. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−93.
28. Clore, G. M., Bax, A., and Gronenborn, A. M. (1991) *J. Biomol. NMR 1*, 13−22.
29. Wüthrich, K., Billeter, M., and Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.
30. Garrett, D. S., Kuszewski, J., Hancock, T. J., Lodi, P. J., Vuister, G. W., Gronenborn, A. M., and Clore, G. M. (1994) *J. Magn. Reson. B 104*, 99−103.
31. Brünger, A. T. (1993) *X-PLOR Manual*, version 3.1, Yale University, New Haven, CT.
32. Wishart, D. S, and Sykes, B. D. (1994) *Methods Enzymol. 239*, 363−392.
33. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Vol. 4, John Wiley and Sons, New York.
34. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1996) *Protein Sci. 5*, 1067−80.
35. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1997) *J. Magn. Reson. 125*, 171−7.
36. Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wüthrich, K. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 13585−90.
37. Cerini, C., Kerjan, P., Astier, M., Gratecos, D., Mirande, M., and Semeriva, M. (1991) *EMBO J. 10*, 4267−77.
38. The *C. elegans* sequencing consortium (1998) *Science 282*, 2012−8.
39. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−38.
40. Predki, P. F., Nayak, L. M., Gottlieb, M. B., and Regan, L. (1995) *Cell 80*, 41−50.
41. Clemons, W. M., Jr., Davies, C., White, S. W., and Ramakrishnan, V. (1998) *Structure 6*, 429−38.
42. Pellecchia, M., Szyperski, T., Wall, D., Georgopoulos, C., and Wuthrich, K. (1996) *J. Mol. Biol. 260*, 236−50.
43. Agalarov, S. C., Sridhar Prasad, G., Funke, P. M., Stout, C. D., and Williamson, J. R. (2000) *Science 288*, 107−13.
44. Nikulin, A., Serganov, A., Ennifar, E., Tishchenko, S., Nevskaya, N., Shepard, W., Portier, C., Garber, M., Ehresmann, B., Ehresmann, C., Nikonov, S., and Dumas, P. (2000) *Nat. Struct. Biol. 7*, 273−7.
45. Escalante, C., and Yang, D. C. (1993) *J. Biol. Chem. 268*, 6014−23.
46. Mirande, M., Lazard, M., Martinez, R., and Latreille, M. T. (1992) *Eur. J. Biochem. 203*, 459−66.
47. Whelihan, E. F., and Schimmel, P. (1997) *EMBO J. 16*, 2968−74.
48. Reed, V. S., Wastney, M. E., and Yang, D. C. (1994) *J. Biol. Chem. 269*, 32932−6.
49. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281−96.
50. Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) *Science 274*, 1531−4.
51. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
52. Merritt, Ethan, A., Bacon, and David, J. (1997) *Methods Enzymol. 277*, 505−524.
53. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−86.
54. Kim, T., Park, S. G., Kim, J. E., Seol, W., Ko, Y.-G., and Kim, S. (2000) *J. Biol. Chem. 275*, 21768−72.

BI001393H