Guest−Host Crosstalk in an Angiogenin−RNase A Chimeric Protein[†,‡]

Daniel E. Holloway,[§] Robert Shapiro,[||] Michelle C. Hares,[§] Demetres D. Leonidas,[§,⊥] and K. Ravi Acharya*[,§]

Department of Biology and Biochemistry, South Building, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom, and Center for Biochemical and Biophysical Sciences and Medicine and Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115

Received May 17, 2002; Revised Manuscript Received June 25, 2002

ABSTRACT: Angiogenin and ribonuclease A share 33% sequence identity but have distinct functions. Angiogenin is a potent inducer of angiogenesis that is only weakly ribonucleolytic, whereas ribonuclease A is a robust ribonuclease that is not angiogenic. A chimera ("ARH-I"), in which angiogenin residues 58−70 are replaced with residues 59−73 of ribonuclease A, has intermediate ribonucleolytic potency and no angiogenic activity. Here we report a crystal structure of ARH-I that reveals the molecular basis for these characteristics. The ribonuclease A-derived (guest) segment adopts a structure largely similar to that in ribonuclease A, and successfully converts this region from a cell-binding site to a purine-binding site. At the same time, its presence causes complex changes in the angiogenin-derived (host) portion that account for much of the increased ribonuclease activity of ARH-I. Guest−host interactions of this type probably occur more generally in protein chimeras, emphasizing the importance of direct structural information for understanding the functional behavior of such molecules.

The proteomes of complex organisms are comprised largely of families and superfamilies. Most, if not all, of the individual members of these groups serve distinct roles, with variations ranging from subtle to extremely dramatic. One approach for studying the molecular basis for this functional diversity has been to generate chimeric proteins in which residues or segments from one homologue are transplanted into another (*1*, *2*). In principle, the properties of hybrids can provide considerable insight into the functional individuality of the two components.

Bovine pancreatic ribonuclease (RNase A,[1] EC 3.1.27.5) and human angiogenin (Ang) provide a particularly intriguing system with which to investigate structure−function relationships among homologues. RNase A is a digestive enzyme with high catalytic efficiency toward general RNA substrates (*3*, *4*). Ang was discovered as a tumor-derived angiogenesis factor and found unexpectedly to be a member of the pancreatic RNase superfamily (*5*−*7*). Although Ang contains the full catalytic apparatus of RNase A, its ribonucleolytic potency is some 4−7 orders of magnitude lower (*1*, *8*). This weak activity is not a mere vestige of evolution: mutational studies have demonstrated that Ang must have an intact enzymatic active site to induce angiogenesis (*9*, *10*). Nonetheless, RNase A and all other superfamily members that have been tested are devoid of angiogenic activity.

Comparison of the sequences of Ang and RNase A reveals that one functionally important segment of RNase A, residues 63−74, is poorly conserved in Ang (Figure 1a): the intrasegment disulfide bridge (65/72) is absent, there is only one apparent sequence identity, and Ang contains two fewer residues overall. Early crystallographic studies (*3*, *4*, *11*) had shown that the RNase segment between Cys 65 and Cys 72 forms one face of the B_2 purine-binding subsite, and might influence the catalytic site as well. To evaluate the significance of the dissimilarity between Ang and RNase A in this region, Harper and Vallee (*1*) prepared a hybrid protein (ARH-I) in which residues 58−70 of Ang were replaced with the corresponding residues, 59−73, of RNase A (Figure 1a). The regional substitution enhanced enzymatic activity by up to several hundred-fold, conferred a more RNase A-like specificity at B_2, and increased the chemical reactivity of catalytic residues. These changes suggested that the RNase A guest segment in ARH-I is well-accommodated into the Ang host and acts much like it does in RNase A. ARH-I was also found to have greatly reduced angiogenic activity (*1*), reflecting the involvement of the substituted segment in the cellular interactions of Ang (*7*, *12*). In a complementary study, Raines et al. reported that the reverse substitution (in which residues 59−73 of RNase A are replaced with residues 58−70 of Ang) confers angiogenic activity upon RNase A and considerably lowers its ribonucleolytic activity (*13*).

The crystal structure of Ang (*14*) confirmed that the three-dimensional (3D) structure of residues 58−70 differs markedly from that of its counterpart in RNase A, but also indicated that the impact of the RNase A-derived segment in ARH-I might be more complex than first envisioned. Specifically, it revealed a potential functional connection in ARH-I, between the substituted segment and the B_1 pyri-

[†] This work was supported by Cancer Research UK Grant SP2354/102 (to K.R.A.), Medical Research Council UK Grant 9540039 (to K.R.A.), and National Institutes of Health Grant CA-88738 (to R.S.).

[‡] The atomic coordinates have been deposited in the Protein Data Bank (entry 1GV7).

* To whom correspondence should be addressed. Telephone: +44-1225-386238. Fax: +44-1225-386779. E-mail: K.R.Acharya@bath.ac.uk.

[§] University of Bath.

[||] Harvard Medical School.

[⊥] Present address: Institute of Biological Research and Biotechnology, The National Hellenic Research Foundation, 48 Vas. Constantinou Ave., Athens 11635, Greece.

[1] Abbreviations: Ang, angiogenin; ARH-I, angiogenin−ribonuclease A hybrid I; RNase A, bovine pancreatic ribonuclease.

10.1021/bi026151r CCC: $22.00
Published on Web 07/24/2002

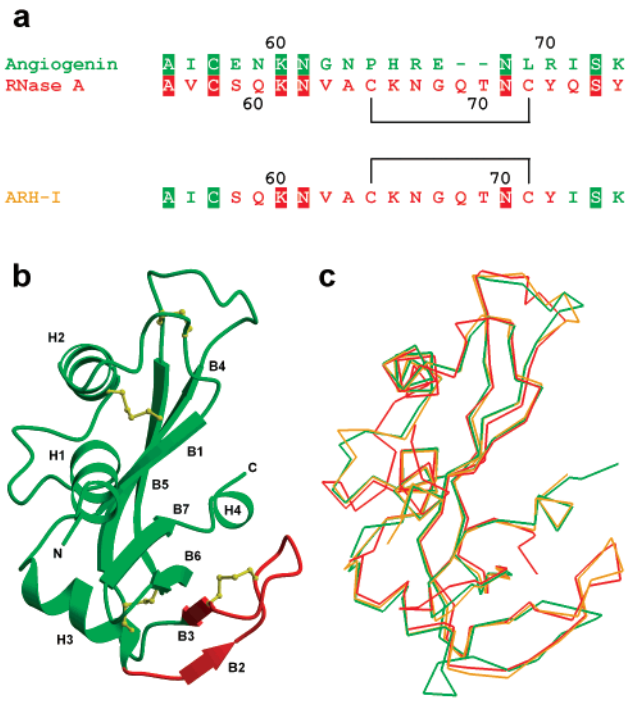

FIGURE 1: Structural overview. (a) Sequences of Ang (green), RNase A (red), and ARH-I (colored according to protein of origin) in the region of mutagenesis. The alignment is that given in ref *1* and does not take the 3D structures into consideration. Residues that appear to be conserved are highlighted, and disulfide-bonded cysteines are bridged. (b) Schematic representation of ARH-I highlighting strands B1−B7 and helices H1−H4. The Ang-derived component is green; the RNase A-derived component is red, and disulfide bridges are yellow. (c) Superposition of the α-carbon traces of ARH-I (gold), Ang (1B1I; green), and RNase A (7RSA; red). The two C-terminal residues of ARH-I are not visible in the crystal structure.

midine-binding site, that does not exist in either parent protein. Further kinetic and mutational studies on ARH-I demonstrated this linkage, while leaving its precise nature unresolved (*15*).

Here we report a crystal structure of ARH-I at 2.1 Å resolution that provides the details of the interactions between the RNase A-derived and Ang-derived regions, and accounts for the characteristic catalytic activity of the hybrid protein.

EXPERIMENTAL PROCEDURES

Protein Purification, Crystallization, and Diffraction Measurements. ARH-I was prepared from the soluble extract of recombinant *Escherichia coli* W3110 by cation-exchange chromatography on SP-Sepharose and Mono-S media as described previously (*1*, *16*), and further purified by C_4 reversed-phase HPLC (*17*). The protein was crystallized by the hanging drop vapor diffusion method. A 4 μL drop (pH 5.8) containing 10 mg/mL protein, 7.5% (w/v) PEG 4000, 4% (v/v) *tert*-butyl alcohol, 0.15 M ammonium acetate, and 0.05 M sodium citrate was equilibrated at 16 °C against a reservoir (0.8 mL) containing 15% (w/v) PEG 4000, 8% (v/v) *tert*-butyl alcohol, 0.3 M ammonium acetate, and 0.1 M sodium citrate at the same pH. Thin, platelike crystals appeared within 3 weeks. Using a single crystal, room-temperature X-ray diffraction data were collected to 2.1 Å resolution on an ADSC Quantum 4R CCD detector at station PX14.1 of the Synchrotron Radiation Source (Daresbury, U.K.). All available data were indexed and integrated using DENZO (*18*), scaled using SCALEPACK (*18*), and converted to amplitudes using TRUNCATE (*19*). The crystal belonged to space group $P2_12_12$ with the following unit cell dimensions: $a = 38.55$ Å, $b = 64.20$ Å, and $c = 61.25$ Å. Detailed data processing statistics are presented in Table 1.

Table 1: X-ray Data Collection and Refinement Statistics

data collection statistics	
resolution (Å)	40−2.1
no. of measured reflections	53460
no. of unique reflections	8426
completeness[a] (%)	90.0 (89.7)
R_{merge}[a,b] (%)	10.8 (58.4)
$I/\sigma(I)$[a]	12.2 (2.0)
refinement statistics	
resolution (Å)	22.9−2.1
R_{cryst}[c] (%)	22.0
R_{free}[d] (%)	27.4
no. of reflections used	8425
no. of protein atoms	1017
no. of citrate atoms	13
no. of water atoms	62
deviations from ideality (rms)	
bond lengths (Å)	0.006
bond angles (deg)	1.3
dihedrals (deg)	24.6
impropers (deg)	0.86
average *B*-factor (Å^2)	
main chain atoms	36.1
side chain atoms	37.4
all protein atoms	36.7
citrate atoms	59.5
water atoms	46.6

[a] Figures in parentheses correspond to the highest-resolution shell (2.18−2.10 Å). [b] $R_{merge} = \sum_h\sum_i|I(h) - I_i(h)|/\sum_h\sum_i I_i(h)$, where $I(h)$ and $I_i(h)$ are the mean and ith measurements of the intensity of reflection h, respectively. [c] $R_{cryst} = \sum_h|F_o - F_c|/\sum_h F_o$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h, respectively. [d] R_{free} is equal to R_{cryst} for a randomly selected 10% subset of reflections not used in the refinement (*45*).

Phase Determination and Refinement. Initial phases were determined by the molecular replacement method (*20*), employing as a search model the coordinates of Ang (PDB entry 1B1I) (*21*) modified by deletion of residues 58−70. Rotation and translation function searches were conducted using data in the range of 25−3 Å, and a clear solution was obtained which, after rigid-body refinement, had a correlation coefficient of 66.2% and an *R*-factor of 41.5%. CNS (*22*) was used to refine the model, adopting a maximum likelihood target for all procedures. Cycles of conjugate gradient minimization, simulated annealing, and geometry minimization were interspersed with calculation of sigmaA-weighted $2F_o - F_c$ and $F_o - F_c$ electron density maps and manual model building using O (*23*). The resolution was increased stepwise, and *B*-factor refinement was carried out once all discernible protein residues had been incorporated. During the final stages of refinement, a citrate ion was introduced into the model and water molecules were added at positions where peaks exceeded 2.5σ in the $F_o - F_c$ electron density map. Only those waters that were positioned at appropriate distances from suitable hydrogen-bonding partners and had temperature factors of <65 Å^2 were kept. The protein residues are well-defined in the final $2F_o - F_c$ electron density map, with the following exceptions: the two C-terminal residues are substantially disordered and are omitted; the segment of residues 86−88 has generally poor density, but has been modeled; and the side chains of Arg 24, Arg

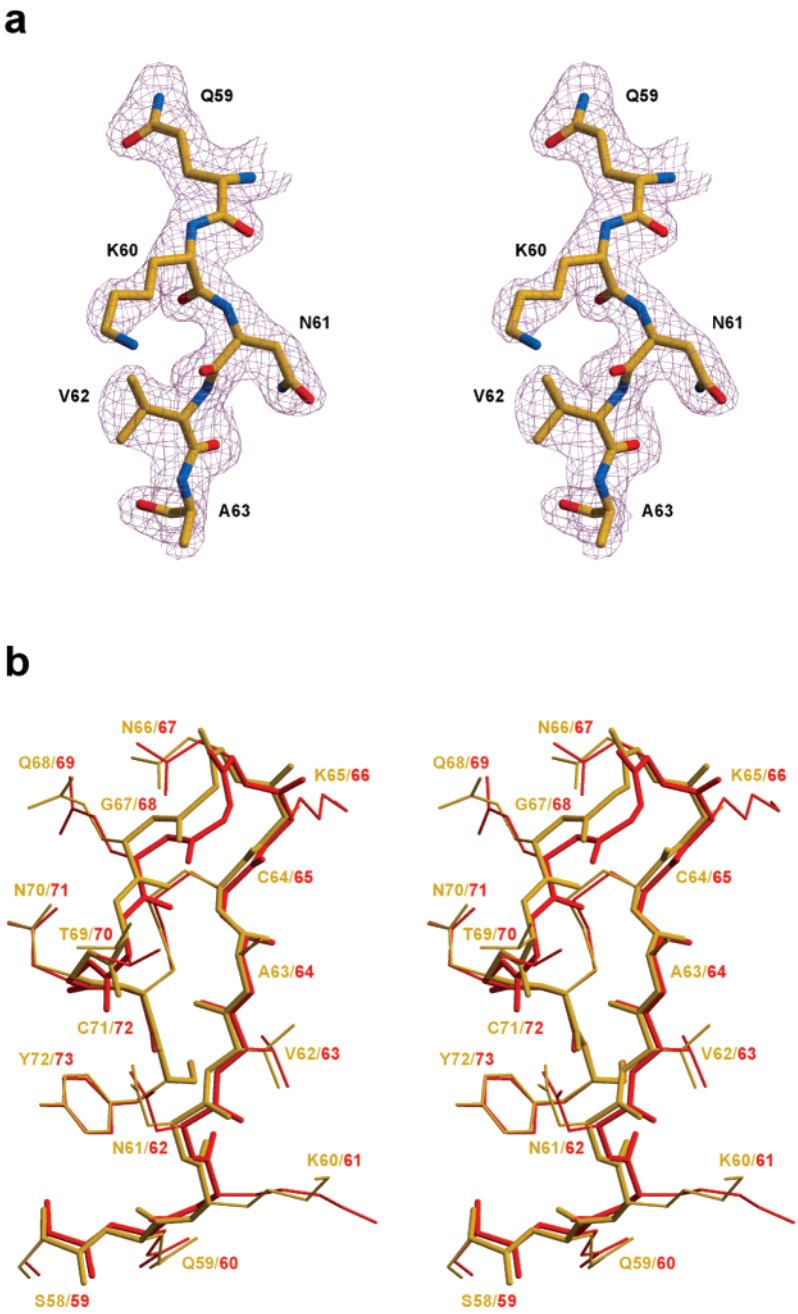

FIGURE 2: RNase A-derived guest segment in ARH-I. (a) Stereoview showing the refined model and $2F_o - F_c$ electron density map (contoured at 1σ) for a representative part of the guest segment. (b) Superposition of the guest segment of ARH-I (gold) and the parental RNase A structure (7RSA; red), shown in stereo. Alternative conformations for RNase A residues Lys 61 and Asn 67 are not shown.

31, Ser 37, Arg 51, Lys 65, and Arg 123 are disordered and have been truncated at C^β. The quality of the final model was assessed with PROCHECK (*24*). Of the 107 discernible non-glycine and non-proline residues, 94 (87.9%) lie in the most favored region of the Ramachandran plot and the others lie in the additional allowed region. Detailed refinement statistics are included in Table 1. Structural superpositions were performed with LSQMAN (*25*). Figures were prepared using BOBSCRIPT (*26*) and RASTER3D (*27*).

RESULTS

Overall Topology. ARH-I adopts the α/β fold common to Ang and RNase A (Figure 1b). The main chain of the Ang-derived portion is not perturbed appreciably by the substitution (Figure 1c). Superposition with the 1.8 Å resolution crystal structure of Ang (PDB entry 1B1I) (*21*) gives an rms deviation of 0.54 Å for 106 C^α atoms (residues 3−57 and 73−123 of ARH-I vs their Ang counterparts). There are only a few short stretches where main chain atoms are >1 Å apart (ARH-I residues 1 and 2, 19 and 20, 87 and 88, and 90 and 91); these deviations are all far from the substituted region and the active site, and in most cases can probably be attributed to differences in the crystal packing arrangements of the two proteins.

The $2F_o - F_c$ map of the guest segment in ARH-I is of high quality (Figure 2a), and reveals a topology strongly resembling that adopted by these residues in RNase A (Figure 1c). Alignment of the 15 C^α atoms of this segment with the corresponding atoms of the 1.26 Å resolution structure of free RNase A (PDB entry 7RSA) (*28*) gives an rms deviation

Table 2: Potential Hydrogen Bonds[a] of the Guest Segment in ARH-I and Their Counterparts in RNase A

ARH-I		RNase A[b]	
bond	length (Å)	bond	length (Å)
Main Chain−Main Chain Interactions			
Ser 58 N−Ala 55 O	3.3	Ser 59 N−Ala 56 O	3.0
Gln 59 N−Ile 56 O	3.0	Gln 60 N−Val 57 O	2.9
Lys 60 N−Ile 73 O	−[c]	Lys 61 N−Gln 74 O	3.3
Val 62 N−Cys 71 O	2.6	Val 63 N−Cys 72 O	3.0
Cys 64 N−Gln 68 O	2.9	Cys 65 N−Gln 69 O	2.8
Gly 67 N−Cys 64 O	3.1	Gly 68 N−Cys 65 O	2.9
Cys 71 N−Val 62 O	3.2	Cys 72 N−Val 63 O	3.4
Tyr 72 N−Val 107 O	2.9	Tyr 73 N−Val 108 O	2.8
Ile 73 N−Lys 60 O	3.2	Gln 74 N−Lys 61 O	2.8
Val 107 N−Tyr 72 O	3.0	Val 108 N−Tyr 73 O	3.0
Cys 109 N−Asn 70 O	3.1	Cys 110 N−Asn 71 O	3.0
Side Chain−Main Chain and Side Chain−Side Chain Interactions			
Ser 58 O^{γ}−Ala 55 O	3.0	Ser 59 O^{γ}−Ala 56 O	−[d]
Asn 61 $N^{\delta 2}$−Thr 69 O	3.1	Asn 62 $N^{\delta 2}$−Thr 70 O	2.8
Asn 61 $N^{\delta 2}$−Cys 71 O	2.7	Asn 62 $N^{\delta 2}$−Cys 72 O	2.9
Lys 65 N−Asp 118 $O^{\delta 2}$	3.2	Lys 66 N−Asp 121 $O^{\delta 2}$	2.8
Lys 65 N^{ζ}−Asp 118 O	−[e]	Lys 66 N^{ζ}−Asp 121 O	3.0
Asn 66 $N^{\delta 2}$−Gln 68 $O^{\epsilon 1}$	−[c]	Asn 67 $N^{\delta 2}$−Gln 69 $O^{\epsilon 1}$	2.9
Gln 68 N−Asn 66 $O^{\delta 1}$	2.8	Gln 69 N−Asn 67 $O^{\delta 1}$	3.0
Gln 68 $N^{\epsilon 2}$−Asn 70 $O^{\delta 1}$	−[c]	Gln 69 $N^{\epsilon 2}$−Asn 71 $O^{\delta 1}$	2.9
Asn 70 $N^{\delta 2}$−Cys 109 O	3.0	Asn 71 $N^{\delta 2}$−Cys 110 O	2.9
Lys 75 N−Gln 59 $O^{\epsilon 1}$	2.5	Tyr 76 N−Gln 60 $O^{\epsilon 1}$	2.9

[a] Hydrogen bonds are listed if the distance between a donor (D) and an acceptor (A) is shorter than 3.3 Å and the D−H−A angle is greater than 90°. [b] PDB entry 7RSA (*28*). [c] Distance of >3.3 Å. [d] Angle of <90°. [e] The side chain of Lys 65 is disordered in the ARH-I crystal.

of 0.38 Å (global alignment of the two molecules gives an rms deviation of 0.66 Å for these atoms). Strands B2 (ARH-I residues 60−63) and B3 (residues 69−73) make three of the four hydrogen bonds present in RNase A (Table 2) to form the same β-hairpin structure. Strand B3 also hydrogen bonds with strand B6 on the main body of the protein (Table 2), as in both RNase A and Ang. Helix H3 (residues 49−59) extends into the substituted segment as in RNase A, rather than terminating with residue 57 as in Ang. Cys 64 and Cys 71, like their RNase A counterparts, form a disulfide bridge. The average main chain temperature factor (37.7 $Å^2$) for this region is similar to that for the rest of the protein (35.9 $Å^2$). This is also the case in RNase A and indicates that the structure adopted by the guest segment is stable.

The preceding observations, taken together, show that ARH-I is an "authentic" hybrid with respect to backbone structure. Consequently, the topological features that define residues 58−70 of Ang as a cell-binding site (*7*) have been lost. The loop of residues 59−61 (H3/B2) of Ang carrying the critical residue Asn 61 has been eliminated, and the loop of residues 66−68 (B2/B3) containing other important cell-binding elements has been lengthened and reoriented.

Side Chain Positions in the Guest Segment. Most of the side chains within the guest segment of ARH-I align reasonably well with their counterparts in RNase A (Figure 2b). The rms deviation for all 49 discernible side chain atoms is 1.08 Å, but reduces to 0.68 Å when Lys 60 and Gln 68, both of which have variable conformations in different RNase A crystal structures, are excluded (the corresponding rms deviations derived from global alignment of the structures are 1.30 and 0.86 Å, respectively). ARH-I residues Ser 58, Asn 61, Ala 63, Cys 64, Asn 70, Cys 71, and Tyr 72 match their RNase A counterparts especially closely. The one notable discrepancy between ARH-I and RNase A in this region involves Lys 65/66. The side chain of Lys 65 is not visible in ARH-I, indicating that it is highly mobile. Although Lys 66 has relatively weak density in some structures of free RNase A and RNase A−nucleotide complexes and is variable in orientation, its conformation is well-defined in each case (*28*−*31*). The N^{ζ} atom of this residue lies within the active site region, usually near the main chain O of Asp 121, with which it hydrogen bonds in some instances (*28*, *29*).

Architecture of the B_2 Site. Crystallographic, NMR, and mutational studies (*3*, *4*, *29*, *30*, *32*, *33*) have shown that the B_2 site of RNase A is formed by the side chains of Cys 65, Asn 67, Gln 69, Asn 71, Ala 109, Glu 111, and His 119. Interactions with the purine base include a functionally important hydrogen bond of Asn 71, transient hydrogen bonds of Asn 67, Gln 69, and Glu111, $\pi-\pi$ stacking with the His 119 imidazole, and van der Waals contacts with Cys 65 and Ala 109. None of the residues from the segment of residues 65−72 of RNase A have structural analogues in Ang, and the Ang residue that corresponds to Glu 111 (Glu 108) is positioned 2−3 Å away from this region. Accordingly, Ang discriminates between bases at B_2 much less effectively; e.g., its preference for A over G is ∼3-fold, compared with values of 12−25-fold for RNase A (*1*, *3*).

The base preference of the B_2 site in ARH-I (29-fold in favor of A) strongly resembles that of RNase A (*1*). The ARH-I crystal structure now indicates that all of the various B_2 components are likely to function similarly (Figure 3). Asn 70 superimposes well with Asn 71 in the RNase A·d(CpA) complex (PDB entry 1RPG) (*29*). The position of Asn 66 corresponds to one of two observed for Asn 67 in 1RPG (lower occupancy) and 7RSA (higher occupancy). The orientation of Gln 68 is different from that of its analogue Gln 69 after C^{γ} because of the proximity of a symmetry-related Trp 91 residue, but Gln 68 should be free to reposition in solution; the conformation seen in ARH-I resembles that in the RNase A·3′-CMP complex (PDB entry 1RPF) (*29*). S^{γ} of Cys 64 aligns well with the equivalent atom of RNase A.

The other residues predicted to comprise the B_2 site of ARH-I originate from the host portion. The C^{β} atoms of ARH-I Ala 108 and RNase A Ala 109 superimpose well. The carboxylate of Glu 110 is almost 2 Å closer to B_2 than is that of Glu 108 in Ang, and its $O^{\epsilon 1}$ atom forms a water-mediated hydrogen bond with Asn 70 $N^{\delta 2}$ which replicates an interaction in 7RSA. His 116 does not stack against a purine ring as does His 119, the equivalent residue, in RNase A complexes with substrate analogues (*29*, *30*); instead, it adopts an alternative "B" conformation seen in some structures of free RNase A and RNase A complexes with pyrimidine nucleotides. In conformation B, the χ_1 and χ_2 torsion angles of the His side chain both differ by ∼160° from those for the productive "A" conformation (which is the conformation of His 114, the equivalent residue, in Ang). The orientation of His 116 in ARH-I appears to be dictated by crystal contacts with Trp 91 from a symmetry-related molecule, and in solution, this side chain would probably be able to play a constructive purine binding role as in RNase A.

Impact of the Guest Segment on the Pyrimidine-Binding Site. In RNase A, residues 59−73 and the residues forming

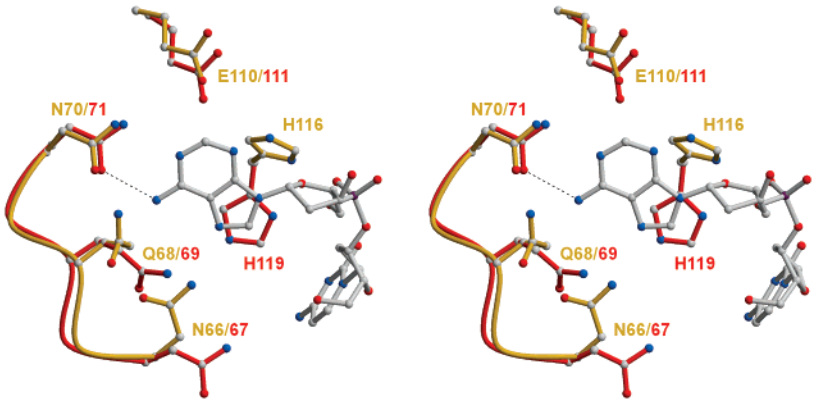

FIGURE 3: Putative B_2 subsite of ARH-I. Stereoview showing key residues involved in the binding of the adenine portion of d(CpA) by RNase A (PDB entry 1RPG) (*29*) superposed with the corresponding region of ARH-I. Residues are shown in ball-and-stick representation. RNase A residues have red sticks; d(CpA) has gray sticks, and ARH-I residues have gold sticks. Balls are colored according to atom type (CPK scheme). For clarity, only side chain atoms are shown, a connecting section of the main chain is drawn as a smooth turn, and the lower-occupancy conformations of RNase A residues Asn 67 and Gln 69 are omitted. The dotted line denotes a key hydrogen bond.

the B_1 pyrimidine site (Thr 45, Val 43, Asn 44, Phe 120, and Ser 123) are separated by ~7 Å at their point of closest approach and appear to be entirely independent structurally. Given the conservation of the B_1 residues of RNase A in Ang (Thr 44, Ile 42, Asn 43, Leu 115, and Ser 118, respectively), it seemed likely prior to the availability of the Ang 3D structure that there would also be no linkage in ARH-I. The Ang crystal structure (*14*) prompted a reassessment of this view by revealing a B_1 site feature unique to Ang that might produce a guest−host clash in ARH-I. Although B_1 in RNase A is fully open to substrate, in Ang it is cryptic: Thr 44, Ile 42, Asn 43, and Leu 115 are all poised to interact with a pyrimidine like their RNase A counterparts, but access to them is blocked by Gln 117. This obstruction reflects the distinctive secondary structure of the C-terminal segment (residues 117−123) of Ang, which places Ser 118 in a position essentially identical to that of Lys 66 in RNase A (*15*) (and ~8 Å from that of its counterpart, Ser 123) (Figure 4a). {Superpositions of Ang with other structures of free (3RN3) (*31*) and complexed RNase A [1RPG (*29*), 1ROB (*29*), 1AFK (*34*), and 1RCN (*30*)] at low or medium ionic strength show similar overlaps.} At this location, Ser 118 forms two hydrogen bonds with Asp 116 (O^{γ}−$O^{\delta 1}$ and NH−$O^{\delta 1}$) that appear to help "push" Gln 117 into B_1, an inference supported by the 7- and 15-fold increases in activity that occur when Ser 118 and Asp 116, respectively, are replaced with Ala (*15, 35*).

Clearly, the ARH-I residues that correspond to Ser 118 of Ang and Lys 66 of RNase A (Ser 120 and Lys 65, respectively) cannot both occupy the same positions as in their proteins of origin. Reorientation of the Ser would likely prevent the replication in ARH-I of the Ser 118−Asp 116 hydrogen bond, and thereby destabilize the native "B_1-closed" conformation or even induce the restructuring proposed to open B_1 during the normal catalytic pathway of Ang (*14*). Alternatively, if the Ser−Lys clash is averted by reorientation of the Lys, the guest segment might not affect the B_1 blockage. Reinvestigation of ARH-I in light of the Ang 3D structure (*15*) showed that the regional replacement has a marked impact on B_1 function, suggesting that the Ser has shifted: (i) ARH-I binds 2′-CMP 10-fold more tightly than does Ang; (ii) ARH-I hydrolyzes cytidine cyclic 2′,3′-phosphate 25-fold more effectively than does Ang; and (iii) replacement of Gln 119 with Ala in ARH-I produces only a minor increase in activity (whereas the corresponding substitution of Gln 117 in Ang enhances activity by 10−14-fold).

Table 3: Potential Hydrogen Bonds[a] of the C-Terminal Regions in ARH-I and Angiogenin

ARH-I		angiogenin[b]	
bond	length (Å)	bond	length (Å)
Main Chain−Main Chain Interactions			
Gln 119 N−Leu 117 O	2.7	Gln 117 N−Leu 115 O	−[d]
Ile 121 N−Asp 118 O	3.3	Ile 119 N−Asp 116 O	−[c]
Phe 122 N−Gln 119 O	2.7	Phe 120 N−Gln 117 O	3.0
Arg 123 N−Ser 120 O	−[c]	Arg 121 N−Ser 118 O	3.1
Side Chain−Main Chain and Side Chain−Side Chain Interactions			
Thr 44 N−Gln 119 $O^{\epsilon 1}$	2.8	Thr 44 N−Gln 117 $O^{\epsilon 1}$	2.9
Gln 119 $N^{\epsilon 2}$−Thr 44 $O^{\gamma 1}$	3.3	Gln 117 $N^{\epsilon 2}$−Thr 44 $O^{\gamma 1}$	3.1
Thr 44 $O^{\gamma 1}$−Thr 82 $O^{\gamma 1}$	3.0	Thr 44 $O^{\gamma 1}$−Thr 80 $O^{\gamma 1}$	2.8
Ser 120 N−Asp 118 $O^{\delta 1}$	−[c]	Ser 118 N−Asp 116 $O^{\delta 1}$	3.1
Ser 120 O^{γ}−Asp 118 $O^{\delta 1}$	−[c]	Ser 118 O^{γ}−Asp 116 $O^{\delta 1}$	2.5
Lys 65 N−Asp 118 $O^{\delta 2}$	3.2		−[e]

[a] Hydrogen bonds are listed if the distance between a donor (D) and an acceptor (A) is shorter than 3.3 Å and the D−H−A angle is greater than 90°. [b] PDB entry 1IBI (*21*). [c] Distance of >3.3 Å. [d] Angle of <90°. [e] Lys 65 has no counterpart in angiogenin.

The ARH-I crystal structure now reveals that in fact **both** of the residues involved in the potential conflict have been repositioned. Ser 120 O^{γ} and NH are shifted by ~3 and 0.9 Å, respectively, vis-à-vis the analogous Ser 118 atoms of Ang, and the hydrogen bonds of these atoms with $O^{\delta 1}$ of Asp 116 in Ang (Asp 118 in ARH-I) have been lost (Figure 4b): distances have been lengthened from 2.5 to 5.1 Å and from 3.1 to 5.6 Å, respectively. As already noted, the location of the Lys 65 side chain cannot be ascertained. Nonetheless, it is clear that this residue cannot be positioned like Lys 66 in RNase A because it would still overlap Ser 120 (N^{ζ} and C^{ϵ} of Lys 66 are both 1.4 Å from C^{β} of Ser 120 in the superimposed RNase A and ARH-I structures).

The perturbation of the C-terminal region of Ang by the guest segment extends beyond the effects on Ser 118/120 (Figure 4b and Table 3). Freed from its hydrogen bonds with this Ser, the side chain of ARH-I Asp 118 lies ~3 Å from the position of Asp 116 in Ang and coincides well with Asp 121 in RNase A. It makes three hydrogen bonds with the main chain of the new segment: one directly with Lys 65

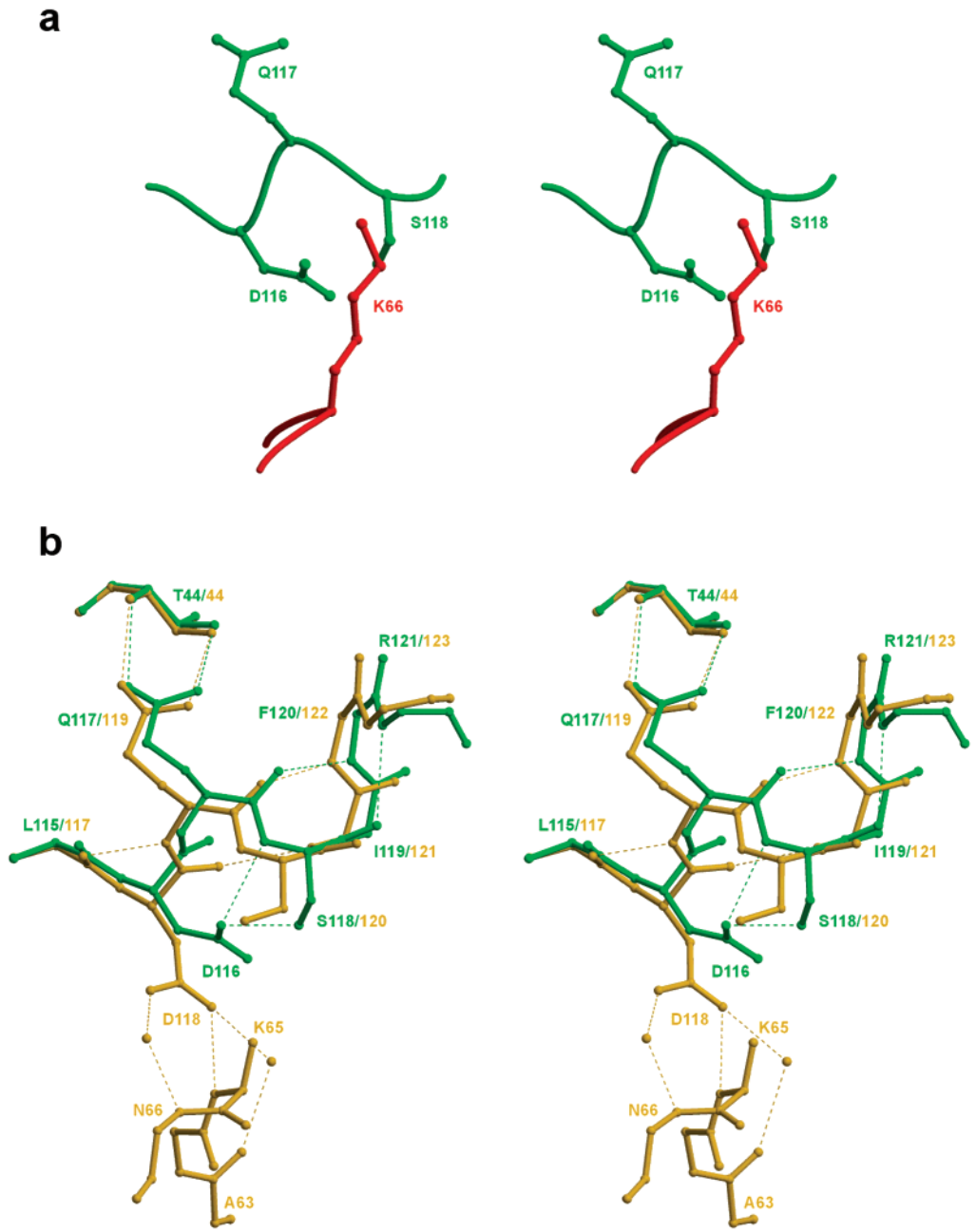

FIGURE 4: Impact of the RNase A-derived guest segment on the C-terminal region of Ang. (a) Stereoview of a superposition of Ang residues 116−119 (1B1I; green) and RNase A residues 65−67 (7RSA; red, only the side chain of Lys 66 is drawn) showing a clash between the positions of Ang Ser 118 and RNase A Lys 66. (b) Stereoview of a superposition of the C-terminal regions of ARH-I (gold) and Ang (green); selected residues that interact with these regions are also shown. The side chain of ARH-I Lys 65 beyond C^{β} is not visible in the crystal structure. The side chains of Ang residues 115 and 119−121, their ARH-I counterparts (residues 117 and 121−123), and ARH-I residues 63, 64, and 66 are omitted for clarity. Dotted lines denote hydrogen bonds, and spheres denote water molecules.

($O^{\delta 2}$−NH) and two via water molecules with Ala 63 ($O^{\delta 2}$−O) and Asn 66 ($O^{\delta 1}$−N). All of these are also present in RNase A crystal structures. In addition, Asp 118 of ARH-I is now well-placed to replicate the hydrogen bond of Asp 121 $O^{\delta 1}$ with $N^{\epsilon 2}$ of His 119 (His 116 in the chimera) when this histidine reorients to its productive conformation (see above). This hydrogen bond in RNase A has been shown to make a significant contribution to catalytic efficiency: replacement of the Asp with Ala diminishes k_{cat}/K_m values by 16−170-fold (*36*). The new set of interactions of Asp 118 in ARH-I is coincident with a minor displacement of the C-terminus toward the β-hairpin [the rms deviation for the C^{α} atoms of residues 118−123 (ARH-I numbering) is 0.71 Å, compared to 0.21 Å for residues 113−117]. The hydrogen bonding arrangement of the C-terminal main chain is somewhat altered in the chimera, but the region retains its 3_{10}-helical conformation.

Other important features of the Ang active site are largely preserved in ARH-I. Gln 119 occupies the same obstructive position as Gln 117 in Ang, forming two comparable hydrogen bonds with Thr 44 (Figure 4b and Table 3); this is consistent with the crystal structure of D116H-Ang (*34*), which shows that removal of the Asp 116−Ser 118 interactions is not sufficient to open the B_1 site. The side chain of the catalytic residue His 13 superimposes almost perfectly in the two structures, as does NH of Leu 115 (Leu 117 in

ARH-I), which is also thought to interact with the scissile phosphodiester (*37*). Two other residues at the catalytic center, Lys 40 and Gln 12, are displaced 0.9−1.5 Å from their counterparts, probably reflecting the formation of hydrogen bonds with a citrate molecule from the crystallization medium.

DISCUSSION

ARH-I was one of the earliest chimeric proteins generated to investigate the structural basis for functional diversity among homologues. At the time, the 3D structure of only one of the parent proteins was known and the properties of the chimera seemed to be fully understandable in terms of simple gain and/or loss of loop function. Now that the structures of the second parent and the chimera itself have been determined, it is clear that that view was inadequate. Thus, although some features of ARH-I—its improved purine binding and lack of angiogenicity—can indeed be explained as originally postulated, others reflect guest−host interactions that could not have been predicted. Most notably, the guest segment induces a partial reorganization of the C-terminal region of the host that accounts for much of the increased enzymatic activity of the chimera. Although Gln 117 (Gln 119 in ARH-I) remains obstructive, the loss of the hydrogen bonds linking Asp 116 (Asp 118) and Ser 118 (Ser 120) should facilitate the opening of the B_1 site to accommodate substrate. Double replacement of both the Asp and Ser with Ala was shown previously to enhance activity toward tRNA by 24-fold (*15*), and it seems likely that at least this much of the ~400-fold overall increase for ARH-I stems from this source. Indeed, the repositioning of Asp 116 (Asp 118) not only eliminates the negative role of this residue but also should allow it to act positively to orient one of the catalytic histidines (residue 116 in ARH-I) as does its RNase A counterpart (*36*). The hydrogen bond of the Asp 118 carboxylate with the main chain N of Lys 65 of the guest segment may be an important factor defining the beneficial location of this Asp. Moreover, it is possible that this interaction, rather than avoidance of the Ser−Lys clash as discussed above, provides the primary driving force for the changes in C-terminal structure.

Guest−host interactions appear to affect the functioning of the guest segment of ARH-I in addition to that of the host. The location of Ser 120 prevents the side chain of Lys 65 from occupying a defined position like its counterpart, Lys 66, in RNase A. The Lys 66 ammonium group has been shown to contribute 0.7 kcal/mol to transition-state binding by forming Coulombic interactions with the phosphoryl group immediately upstream from the scissile bond (*38*). It seems unlikely that Lys 65 in ARH-I would make a comparable contribution, given its extreme flexibility.

Transplantations of segments, residues, or entire domains have been used to investigate functional differences between many other pairs of related proteins (*2*), as well as to generate proteins with novel activities (*39*), to eliminate undesirable properties such as antigenicity [as with humanized monoclonal antibodies (*40*)], and to address basic questions in protein evolution (*41*). Our search of the Protein Data Bank yielded no more than a dozen instances in which structures of a chimera and both of its parents have been deposited, and in only a few of them have the replacements involved relatively small structural units (e.g., β-hairpins, turns, or loops) as for ARH-I. Effects of guest on host were reported for a hybrid of staphylococcal nuclease and concanavalin A (*42*), and of host on guest for a charybdotoxin−curaremimetic toxin chimera (*43*). In the first case, the loss of contacts between the retained and replaced regions of the host significantly reduced thermodynamic stability. In the second, the low affinity for binding of the chimera to a receptor was proposed to reflect, in part, host-imposed deviations of the guest segment from the structure adopted in its parent. However, the authors of this report noted that some residues involved in receptor binding had not been transplanted, and suggested that this might also account for the deficient binding.

To our knowledge, ARH-I is the first chimeric protein for which it has been possible to explain complex functional characteristics in terms of guest−host interactions. However, it seems likely that this type of crosstalk is not uncommon, even where chimera design has aimed to avoid it. Loops and other structural units targeted for transplantation may often be (or appear to be) independent in both parents while impacting on each other heavily in the chimera constructed from them. Similarly, the natural evolution of novel protein functions through the acquisition of smaller mutations can, in some cases, be attributed not to localized perturbations but to a change in the interplay between distinct sites (*44*). Therefore, considerable caution should be exercised when interpreting the functional properties of chimeras in the absence of detailed structural information about all three proteins.

ACKNOWLEDGMENT

We thank members of the Structural Biology Group at Bath and J. F. Riordan at Harvard Medical School for constructive criticism of the manuscript.

REFERENCES

1. Harper, J. W., and Vallee, B. L. (1989) *Biochemistry 28*, 1875−1884.
2. Cunningham, B. C., Jhurani, P., Ng, P., and Wells, J. A. (1989) *Science 243*, 1330−1336.
3. Richards, F. M., and Wyckoff, H. W. (1971) in *The Enzymes* (Boyer, P. D., Ed.) pp 647−806, Academic Press, New York.
4. Raines, R. T. (1998) *Chem. Rev. 98*, 1045−1065.
5. Fett, J. W., Strydom, D. J., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., and Vallee, B. L. (1985) *Biochemistry 24*, 5480−5486.
6. Strydom, D. J., Fett, J. W., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., and Vallee, B. L. (1985) *Biochemistry 24*, 5486−5494.
7. Riordan, J. F. (1997) in *Ribonucleases: Structures and Functions* (D'Alessio, G., and Riordan, J. F., Eds.) pp 445−489, Academic Press, New York.
8. Shapiro, R., Riordan, J. F., and Vallee, B. L. (1986) *Biochemistry 25*, 3527−3532.
9. Shapiro, R., Fox, E. A., and Riordan, J. F. (1989) *Biochemistry 28*, 1726−1732.
10. Shapiro, R., and Vallee, B. L. (1989) *Biochemistry 28*, 7401−7408.
11. Wodak, S. Y., Liu, M. Y., and Wyckoff, H. W. (1977) *J. Mol. Biol. 116*, 855−875.
12. Hallahan, T. W., Shapiro, R., and Vallee, B. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2222−2226.
13. Raines, R. T., Toscano, M. P., Nierengarten, D. M., Ha, J. H., and Auerbach, R. (1995) *J. Biol. Chem. 270*, 17180−17184.
14. Acharya, K. R., Shapiro, R., Allen, S. C., Riordan, J. F., and Vallee, B. L. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2915−2919.
15. Shapiro, R. (1998) *Biochemistry 37*, 6847−6856.
16. Shapiro, R., and Vallee, B. L. (1992) *Biochemistry 31*, 12477−12485.

17. Holloway, D. E., Hares, M. C., Shapiro, R., Subramanian, V., and Acharya, K. R. (2001) *Protein Expression Purif. 22*, 307−317.
18. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
19. Bailey, S. (1994) *Acta Crystallogr. D50*, 760−763.
20. Navaza, J. (1994) *Acta Crystallogr. D50*, 157−163.
21. Leonidas, D. D., Shapiro, R., Allen, S. C., Subbarao, G. V., Veluraja, K., and Acharya, K. R. (1999) *J. Mol. Biol. 285*, 1209−1233.
22. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., GrosseKunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
23. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 109−110.
24. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
25. Kleywegt, G. J., and Jones, T. A. (1994) *CCP4/ESF-EACBM Newsletter on Protein Crystallography 31*, 9−14.
26. Esnouf, R. M. (1997) *J. Mol. Graphics 15*, 132−134.
27. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
28. Wlodawer, A., Svensson, L. A., Sjölin, L., and Gilliland, G. L. (1988) *Biochemistry 27*, 2705−2717.
29. Zegers, I., Maes, D., Dao-Thi, M.-H., Poortmans, F., Palmer, R., and Wyns, L. (1994) *Protein Sci. 3*, 2322−2339.
30. Fontecilla-Camps, J. C., de Llorens, R., le Du, M. H., and Cuchillo, C. M. (1994) *J. Biol. Chem. 269*, 21526−21531.
31. Howlin, B., Moss, D. S., and Harris, G. W. (1989) *Acta Crystallogr. A45*, 851−861.
32. Toiron, C., Gonzales, C., Bruix, M., and Rico, M. (1996) *Protein Sci. 5*, 1633−1647.
33. Tarragona-Fiol, A., Eggelte, H. J., Harbron, S., Sanchez, E., Taylorson, C. J., Ward, J. M., and Rabin, B. R. (1993) *Protein Eng. 6*, 901−906.
34. Leonidas, D. D., Shapiro, R., Subbarao, G. V., Russo, A., and Acharya, K. R. (2002) *Biochemistry 41*, 2552−2562.
35. Harper, J. W., and Vallee, B. L. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7139−7143.
36. Schultz, L. W., Quirk, D. J., and Raines, R. T. (1998) *Biochemistry 37*, 8886−8898.
37. Leonidas, D. D., Chavali, G. B., Jardine, A. M., Li, S., Shapiro, R., and Acharya, K. R. (2001) *Protein Sci. 10*, 1669−1676.
38. Fisher, B. M., Ha, J.-H., and Raines, R. T. (1998) *Biochemistry 37*, 12121−12132.
39. Vita, C. (1997) *Curr. Opin. Biotechnol. 8*, 429−434.
40. Verhoeyen, M., Milstein, M., and Winter, G. (1988) *Science 239*, 1534−1536.
41. Hopfner, K.-P., Kopetzki, E., Kresse, G.-B., Bode, W., Huber, R., and Engh, R. A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 9813−9818.
42. Hynes, T. R., Kautz, R. A., Goodman, M. A., Gill, J. F., and Fox, R. O. (1989) *Nature 339*, 73−76.
43. Zinn-Justin, S., Guenneugues, M., Drakopoulou, E., Gilquin, B., Vita, C., and Ménez, A. (1996) *Biochemistry 35*, 8535−8543.
44. Zhang, J., and Rosenberg, H. F. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 5486−5491.
45. Brünger, A. T. (1992) *Nature 355*, 472−474.

BI026151R