Crystallographic structure of an active, sequence-engineered ribonuclease

(synthetic modeling/protein engineering/semisynthesis/x-ray diffraction/peptide–protein recognition)

HOPE C. TAYLOR*, AKIRA KOMORIYA†, and IRWIN M. CHAIKEN‡§

*Department of Biochemistry, Duke University Medical Center, Durham, NC 27710; †Revlon Biotechnology Research Center, 2501 Research Boulevard, Rockville, MD 20850; and ‡Molecular, Cellular and Nutritional Endocrinology Branch, National Institute of Arthritis, Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20205

Communicated by Christian B. Anfinsen, June 3, 1985

ABSTRACT X-ray diffraction methods were used to test a synthetic-modeling approach to the sequence engineering of bovine pancreatic ribonuclease. A model of RNase S-peptide (residues 1–20), having a simplified amino acid sequence but retaining elements deduced to be essential for conformation and function, was previously synthesized and found to form a catalytically active and stable complex with native S-protein (residues 21–24). We have now obtained a 3-Å-resolution electron density map of this semisynthetic complex which reveals that the conformation of model peptide closely mimics that of native S-peptide, as intended by sequence design. Some small differences from the native structure are observed: Glu-2 and Arg-10 of the model complex are not close enough to form a salt bridge, the position of the His-12 imidazole ring is slightly shifted in the active site, and the peptide’s amino terminus is reoriented. Nonetheless, the major structural features predicted to be essential by computer-aided peptide-design analysis are preserved in the model peptide portion of the complex. These include (i) the α-helical framework involving residues 3–13, (ii) the catalytically competent orientation of His-12, and (iii) complex-stabilizing nonbonding interactions involving Phe-8 and Met-13 of S-peptide and hydrophobic residues in the cleft region of S-protein. Further, sequence simplification has not introduced any non-native, potentially stabilizing contacts between the model peptide and S-protein. The results emphasize the usefulness, in redesigning native proteins, of categorizing sequence into residues providing conformational framework and those determining intra- and intermolecular surface recognition.

Increasing understanding of the principles by which amino acid sequences govern protein structure (1–5) provides the opportunity to engineer proteins with altered physical and functional properties by designing new or modified sequences. Such engineered proteins can be obtained both by redesigning native proteins (e.g., by site-specific, including multisite, mutagenesis) and by fashioning altogether new sequences. Chemical and enzymatic synthesis (6–8) as well as recombinant DNA methods (9–11) can be used to produce peptides and proteins of planned non-native sequence, with the method of choice depending on the size and type of sequence desired. However, defining a precise sequence to obtain a particular conformation and function has been hindered by an incomplete set of predictive rules relating sequence and higher-order molecular properties. The use of semisynthesis allows sequence engineering of one portion of a protein that can be recombined noncovalently or covalently with the remainder of the protein in its native form. Both theoretical sequence design and experimental synthesis are thus simplified.

The general problem of reengineering native proteins has been addressed previously (12, 13) by defining some simplified design guidelines with the test system bovine pancreatic RNase S [the fully active noncovalent fragment complex obtained from RNase A by limited subtilisin-catalyzed proteolysis and containing S-peptide (residues 1–20) and S-protein (residues 21–124) (14)]. We have sought to categorize the structural features most essential for RNase S-peptide to produce a functional, noncovalent complex with S-protein and have sought to demonstrate the critical role of these deduced factors through design, synthesis, and characterization of model peptides. Our approach to engineering such peptides assumes that, to a first approximation, complex conformational and functional information encoded in the amino acid sequence can be sorted into secondary and tertiary structural elements that include intra- and intermolecular recognition surfaces and catalytic sites.

A computer-assisted structural analysis of RNase S-peptide was reported (15) based on the native RNase S crystal structure and a local-interaction analysis of residues involved in the nonbonding interactions between S-peptide and S-protein. This consisted of identification and evaluation of the free energy contribution of each side-chain contact present between native S-peptide and S-protein to the total complex-stabilizing side-chain–side-chain contact free energy. The magnitude of the van der Waals stabilization free energy associated with a specific side-chain contact was assessed on a pairwise basis by means of a set of interaction parameters that resembles, in its formalism, Hildebrand’s solubility parameters (16). Each interaction parameter is characteristic of a single type of side chain. These theoretical analyses emphasized the importance of (i) nonbonding interactions involving Phe-8 and Met-13, which contribute a major complex-stabilizing free energy, and (ii) the α-helical framework of the peptide in orienting these two side chains optimally toward hydrophobic residues Val-47, -54, and -108 and Leu-51 of S-protein.

Previous synthetic modification studies involving single-residue substitutions showed the importance of helix-forming potential in S-peptide for stabilization of RNase S complex as well as the catalytic role of His-12 (12, 17, 18). Further, residues 16–20 of the native S-peptide were found to be considerably disordered in the native RNase S crystal structure (19) and unnecessary for complex formation, catalytic activity, and ordered structure (ref. 7 and references therein). Recent crystallographic studies of the analogue complex of native S-protein and a residue 1–15 synthetic peptide containing 4-fluorohistidine at residue 12 (20) showed no change relative to the native peptide in either overall conformation or the orientation of specific side chains in the active-site region.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked ‘advertisement’ in accordance with 18 U.S.C. §1734 solely to indicate this fact.

§To whom reprint requests should be addressed.
pared

bone
structure
KY424, CA

ribbon is shown truncated conserved residues (differences from atations), Richardson.
Phe-8 and Met-13.

pentadecapeptide with native to as to design of interaction properties ical designed the model semisynthetic we

S-peptide. As catalytically effective of native structure formation of complex a and function in the model semisynthetic structure to native RNase S, and previously crystallized semisynthetic RNase S analogues (20–22). Three angstrom-resolution x-ray-diffraction data were collected from a single crystal, using nickel-filtered Cu Kα radiation. After corrections for radiation-induced intensity decay, background scattering, and absorption, data were merged with phase information calculated from modified coordinates of refined RNase S (20); i.e., atoms were deleted from the structure-factor calculation as necessary to account for sequence simplification. These same modified RNase S coordinates served as a starting molecular model for fitting to electron density maps on the GRIP-75 molecular graphics system of the University of North Carolina Department of Computer Science. “Fragment ΔF” difference maps, obtained using the observed structure amplitudes and the results of a structure-factor calculation from which coordinates of specific atoms of interest had been removed (for a Fourier synthesis of the type \( F_{\text{obs}} - F_{\text{calc}} \), \( \phi_{\text{calc}} \) were used for initial fitting. Our final model-peptide coordinates are the result of fitting model-peptide structure to \( 2(F_{\text{obs}}) - F_{\text{calc}} \) density maps followed by applying stereochemical constraints to the fitted structure with the MODELFIT algorithm (23). There being no evidence for position shifts in the S-protein portion of the maps relative to native RNase S, all fitting was confined to the model peptide.

RESULTS AND DISCUSSION
A comparison of the backbone of the model semisynthetic complex with that of native RNase S is given as superim-

His-12, Phe-8, and Met-13 side chains; investigating the role of other conserved non-alanyl residues; and detecting any compensating nonbonding interactions that might have resulted from sequence simplification.

EXPERIMENTAL PROCEDURES
Crystals of the model RNase S complex were grown from ammonium sulfate/cesium chloride solution buffered with 0.1 M sodium acetate (pH 5.3), as reported (11). Crystals were mounted with a small amount of mather liquor in glass capillaries, which then were sealed. Examination of zero-level precession photographs revealed the crystals to be of the same space group (P3121) and approximate unit-cell dimensions \( (a = b = 44.4 \ \text{Å}, c = 97.2 \ \text{Å}) \) as the “Y” form of native RNase S (19) and previously crystallized semisynthetic RNase S analogues (20–22). Three angstrom-resolution x-ray-diffraction data were collected from a single crystal, using nickel-filtered Cu Kα radiation. After corrections for radiation-induced intensity decay, background scattering, and absorption, data were merged with phase information calculated from modified coordinates of refined RNase S (20); i.e., atoms were deleted from the structure-factor calculation as necessary to account for sequence simplification. These same modified RNase S coordinates served as a starting molecular model for fitting to electron density maps on the GRIP-75 molecular graphics system of the University of North Carolina Department of Computer Science. “Fragment ΔF” difference maps, obtained using the observed structure amplitudes and the results of a structure-factor calculation from which coordinates of specific atoms of interest had been removed (for a Fourier synthesis of the type \( F_{\text{obs}} - F_{\text{calc}} \), \( \phi_{\text{calc}} \) were used for initial fitting. Our final model-peptide coordinates are the result of fitting model-peptide structure to \( 2(F_{\text{obs}}) - F_{\text{calc}} \) density maps followed by applying stereochemical constraints to the fitted structure with the MODELFIT algorithm (23). There being no evidence for position shifts in the S-protein portion of the maps relative to native RNase S, all fitting was confined to the model peptide.

RESULTS AND DISCUSSION
A comparison of the backbone of the model semisynthetic complex with that of native RNase S is given as superim-

Fig. 1. A comparison of the crystallographically determined backbone structure (3 Å resolution) of the model semisynthetic RNase S analogue complex (solid ribbons) and native RNase S (differences from the model are shown as dotted ribbons). The native peptide ribbon is shown truncated at residue 15, as residues 16–20 are disordered in the crystal structure. Drawing courtesy J. S. Richardson.

Fig. 2. A stereo view of the model peptide (stick figure in gold, all non-hydrogen atoms) and the solvent-accessible Connolly dot surface (23) of S-protein [surface associated with all hydrophobic residues (Ala, Val, Leu, Ile, Pro, Phe, and Met) is colored gold, all others green]. On the left image, residue numbers are denoted adjacent to the side chains for Ala-1, Glu-2, Ala-3, Lys-7, Phe-8, Arg-10, His-12, Met-13, and Ala-15.
posed α-carbon ribbon diagrams in Fig. 1. Circular-dichroism spectra of the model peptide in solution and its estimated free energy of α-helical stabilization had suggested at least as great a tendency for this peptide to form α-helical structure as for the native S-peptide (12). Consistent with the intended design, the model peptide is in a predominantly α-helical configuration in the cleft of S-protein, though its amino terminus has rotated significantly relative to the native coordinates.

The structure of the model peptide (all non-hydrogen atoms) in model semisynthetic RNase S is shown in Fig. 2 as a stick diagram superimposed on the solvent-accessible van der Waal’s surface of S-protein generated by the MS algorithm of Connolly (24). Dots associated with the surface of the most hydrophobic amino acids have been colored gold to highlight residues potentially involved in nonbonding interactions with the peptide. Val-47, -54, and -108, Leu-51, and Phe-120 form the hydrophobic surface in the immediate vicinity of Phe-8 and Met-13. The Phe-8 side chain of the model peptide has rotated almost 90° around the bond between α and β carbons relative to its position in native S-peptide (Fig. 3), but the aromatic ring is still as deeply buried and well-packed against the hydrophobic surface.

Met-13 is essentially unmoved from its native position. Nonbonding interactions critical to complex formation have therefore been preserved. Fig. 4 shows the extent to which Met-13 and Phe-8 in the model peptide protrude into the hydrophobic region involved on the S-protein surface. Lee and Richards (25) have shown, in their surface-accessibility studies of RNase S, that the surface of the cleft region consists of similar numbers of polar and nonpolar residues. As emphasized in Figs. 2–4, however, the localized distribution of residues and the topography of surfaces defined by them appear to be most important in optimizing complex-stabilizing free energy derived from hydrophobic interactions. No residues other than Phe-8 and Met-13 in the simplified peptide are found to be involved in compensating nonbonding interactions that could stabilize complex formation with S-protein.

In the native RNase S crystal structure, the proximity of side chains of Glu-2 and Arg-10 had been interpreted to indicate a possible salt-bridge interaction (12). The side chain of Glu-2 in the model peptide, however, points almost directly away from Arg-10, rendering any interaction between the two highly unlikely. Further, no other salt bridges, a feature considered as a possible contributor to the folding of S-peptide alone (26), are found in the complexed model peptide. This result casts doubt on an obligatory role for such a salt bridge in the conformational stabilization of native RNase S.

Proposed mechanisms for RNase A (and S) enzymatic activity call for the involvement of His-12 and -119 as alternating proton donor and acceptor in transphosphorylation and hydrolysis of substrate (27). The δ-nitrogens of these two histidyl imidazole rings are 10 Å apart in the model complex, ~1 Å further than in the native complex. This increased distance may affect the catalytic competence of the active center and thus account for the lowered specific activity of the model RNase S complex (12).

The deductions resulting from earlier computer analysis and solution studies which led to the design of the synthetic model peptide have been confirmed by the 3 Å resolution structure of its analogue complex with native S-protein. This model semisynthetic system offers opportunities for further structural and dynamic calculations, followed by increasingly refined synthetic modeling as a means to evaluate the nature and magnitudes of various molecular forces that govern protein structure and function.

Crystallography was performed in the laboratory of David and Jane Richardson at Duke University. We extend our thanks for their
help in all aspects of this work and for the support of their National Institutes of Health Grant GM15000 and Contract 26300010617-00. The staff of the Molecular Graphics Project of the University of North Carolina Department of Computer Science (supported by National Institutes of Health Grants RR00898 and RR02170), especially Doris Knecht, deserve our gratitude for their patient and expert assistance in visualizing the model peptide’s structure and interactions.