Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 1880–1884, March 1992
Biochemistry

Functional assembly of a randomly cleaved protein

(split proteins/protein assembly/evolution/fragment complementation/aminoacyl-tRNA synthetase)

KIYOTAKA SHIBA AND PAUL SCHIMMEL

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Paul Schimmel, November 18, 1991

ABSTRACT **The sequence of a 939-amino acid polypeptide that is a member of the aminoacyl-tRNA synthetase class of enzymes has been aligned with sequences of 15 related proteins. This alignment guided the design of 18 fragment pairs that were tested for internal sequence complementarity by reconstitution of enzyme activity. Reconstitution was achieved with fragments that divide the protein at both nonconserved and conserved sequences, including locations proximal to or within elements believed to form critical elements of secondary structure. Structure assembly is sufficiently flexible to accommodate fusion of short segments of unrelated sequences at fragment junctions. Complementary chain packing interactions and chain flexibility appear to be widely distributed throughout the sequence and are sufficient to reconstruct large three-dimensional structures from an array of disconnected pieces. The results may have implications for the evolution and assembly of large proteins.**

Limited proteolytic cleavage of ribonuclease A, staphylococcal nuclease, cytochrome *c*, cytochrome b_5 reductase, human pituitary growth hormone, human hemoglobin, thioredoxin C, adenylate kinase, or alanine racemase yields fragments that can reassociate *in vitro* to reconstitute activity (1–10). These protease-generated fragments are believed to represent well-defined structural units or domains that associate through complementary sequences at the interfaces. While domains cannot always be released by proteolysis, when produced by genetic methods they still can participate in noncovalent assembly of protein structure (11–13). Reconstitution of protein structure has also been achieved with overlapping fragment pairs of staphyloccocal nuclease, cytochrome *c*, and SecA, where overlaps of 10 or more amino acids have been used (14–17). In a related vein, a specific protein fragment derived from dihydrofolate reductase (DHFR) can strongly inhibit the refolding of DHFR (18). Inhibition is believed to be caused by binding of the fragment to its complementary site on DHFR, and displacement of the corresponding intrachain segment within the protein.

To investigate more generally whether complementary chain packing interactions are sufficiently strong to overcome breaks in the covalent structure and to displace steric obstacles introduced at the break points, we chose *Escherichia coli* isoleucyl-tRNA synthetase. This monomeric protein of 939 amino acids is one of ten class I aminoacyl-tRNA synthetases (19) and is believed to be historically related to a subgroup that includes cysteinyl-, methionyl-, leucyl- and valyl-tRNA synthetases (20). Although these *E. coli* proteins vary in size from 461 (cysteine enzyme) to 951 (valine enzyme) amino acids, they share sequence motifs in their amino-terminal halves that are part of a Rossmann (nucleotide-binding) fold. For one member of this subgroup of enzymes—*E. coli* methionyl-tRNA synthetase—a three-dimensional structure of the amino-terminal 547 amino acids has been described (21). Structural modeling and mutagenesis suggest that the isoleucine enzyme is folded like methionyl-tRNA synthetase (20, 22–24), so that the interpretation of the functional assembly of "randomly" cleaved isoleucine tRNA synthetase can reasonably be done in the light of a structural model.

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.

MATERIALS AND METHODS

Computer Programs and Sources of Sequences. GAP, PILEUP, PROFILE, and PRETTY programs are provided in the Genetics Computer Group (Madison, WI) software. GAP uses the alignment method of Needleman and Wunsch (25). PILEUP uses the progressive alignment method of Feng and Doolittle (26). PROFILE is based on the methods of Gribskov *et al.* (27).

Synthetase sequences were as follows: *E. coli* isoleucine (19), *Saccharomyces cerevisiae* isoleucine (28), *Methanobacterium thermoautotrophicum* isoleucine (29), *E. coli* valine (30), *S. cerevisiae* valine (cytoplasmic and mitochondrial) (31), *Bacillus stearothermophilus* valine (32), *S. cerevisiae* mitochondrial leucine (33), *Neuropora crassa* mitochondrial leucine (34), *N. crassa* leucine (35), *E. coli* methionine (36), *Thermus thermophilus* methionine (37), *S. cerevisiae* mitochondrial methionine (38), *S. cerevisiae* methionine (39), and *E. coli* cysteine (20, 40).

Plasmid Constructions and Strains. Plasmid pKS148 was assembled from the following DNA fragments: *tet* gene and p15A *ori* of pACYC184 (41), f1 *ori* of pTZ19U (42), and *lacPO–lacZ'* of pTZ19R (42), which contains multiple cloning sites. The universal translation terminator sequence, 5'-GCTTAATTAATTAAGC-3' (Pharmacia), was inserted at the multiple cloning sites. Strain IQ843/pRMS711 is a derivative of *E. coli* K-12 strain MC4100 (43) that contains *ΔileS203::kan* and *recA56* (unpublished work). Viability of this strain is maintained by pRMS711 (R. Starzyk and P.S., unpublished work), which is a derivative of the temperature-sensitive plasmid pMT101 (44) and has a DNA fragment encoding the intact *ileS–lsp* genes.

Biochemical Procedures. To investigate the activity of cell extracts, the 100,000 × *g* supernatant (S100) fractions of complemented cells were prepared and applied to a Superose 12 column (Pharmacia) with 50 mM potassium phosphate buffer, pH 7.5/0.1 mM NaCl/50 mM 2-mercaptoethanol as elution buffer (flow rate, 0.3 ml/min). Each fraction (0.3 ml) was assayed for aminoacylation activity (23).

For Western blot analysis, cells were boiled in SDS sample buffer and fractionated by electrophoresis through an SDS/10% polyacrylamide gel. Proteins were transferred onto an Immobilon-P membrane (Millipore), and enzyme fragments were detected by using anti-isoleucyl-tRNA synthetase serum (23) and the ECL detection system (Amersham).

RESULTS AND DISCUSSION

Multiple Sequence Alignment. There are 15 amino acid sequences available for the subgroup that constitutes the

aforementioned five class I aminoacyl-tRNA synthetases. These sequences were aligned with the GAP, PILEUP, and PROFILE programs (25–27), with conserved amino acids identified with PRETTY (Fig. 1). Although these enzymes vary in size from 461 (*E. coli* cysteinyl-tRNA synthetase) to 1123 (*N. crassa* cytoplasmic leucyl-tRNA synthetase) amino acids, their sequences can be aligned when variable-length insertions are taken into account (20, 22). (These insertions are highlighted in black in Fig. 1.) Structural motifs in *E. coli* methionyl-tRNA synthetase can be placed alongside this alignment and include the amino-terminal nucleotide-binding fold (first 361 amino acids of methionyl-tRNA synthetase) of alternating β-strands and α-helices. This structure contains the sites for binding ATP and amino acid and for aminoacyl adenylate synthesis. A common feature of the nucleotide-binding fold of class I aminoacyl-tRNA synthetases is the

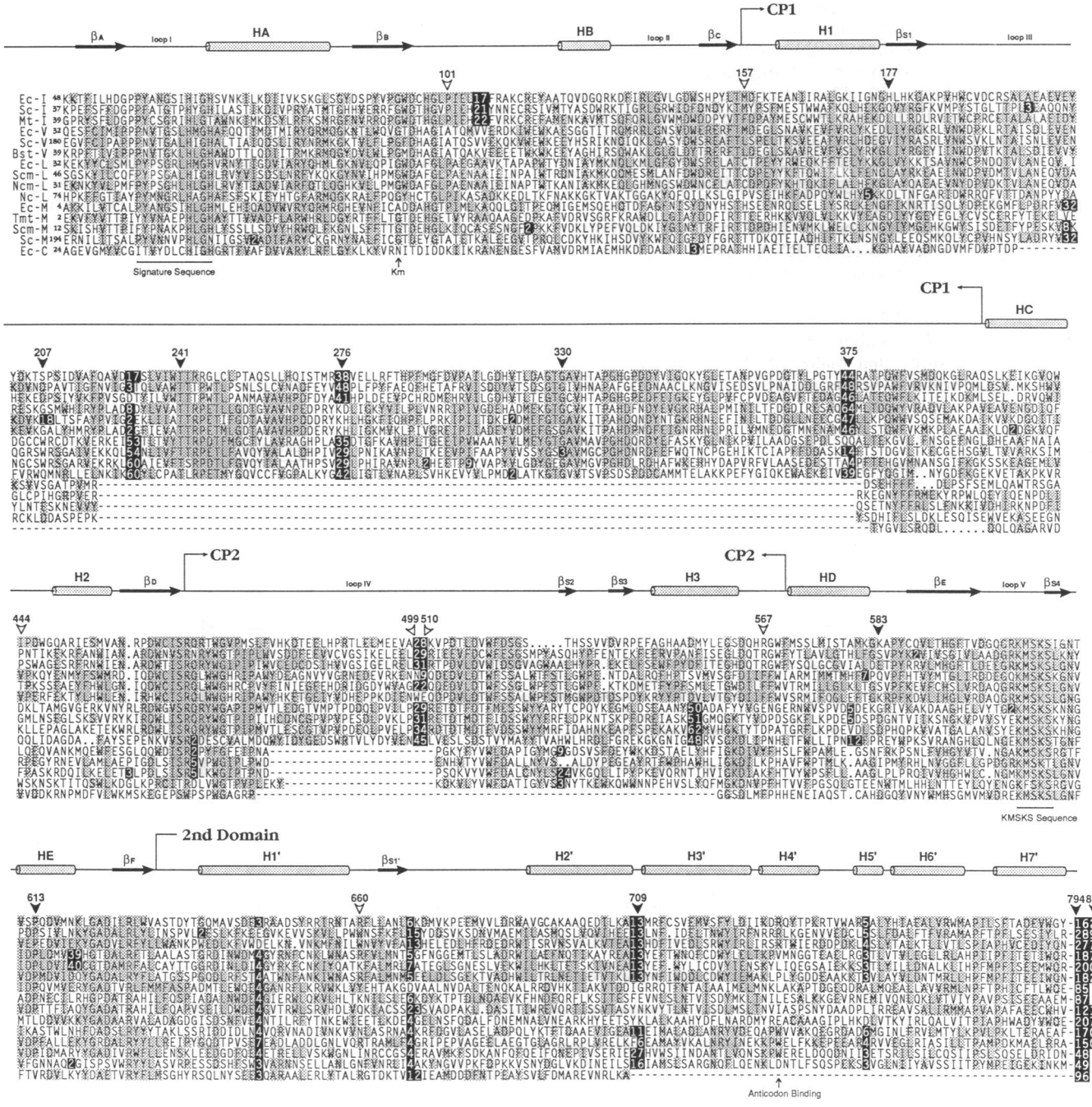

FIG. 1. Alignment of the sequences of isoleucyl-, valyl-, leucyl-, methionyl- and cysteinyl-tRNA synthetases. Conserved residues that were calculated by the PRETTY program are shaded, while dots and hyphens indicate gaps. Numbers in black boxes indicate insertions that could not be aligned, and include the carboxyl-terminal region. The secondary structure elements of *E. coli* methionyl-tRNA synthetase are shown across the top. Nomenclatures of β-strands (beta), α-helices (H), and loops are based on ref. 21. The strands β_A through β_F and the helices HA through HE constitute the Rossmann (nucleotide-binding) fold. CP, connective polypeptide. The conserved signature sequence (ending in the tetrapeptide HIGH) and the KMSKS pentapeptide are indicated on the bottom. Km is the location of $Gly^{94} \rightarrow Arg$ mutation (in Ec-I), which elevates the K_m for isoleucine (23). The proposed anticodon recognition site (45) is also shown. Residues at which *E. coli* isoleucyl-tRNA synthetase were split are indicated by filled arrowheads (split enzymes that reconstitute activity) or open arrowheads (do not reconstitute activity). Ec, *E. coli*; Sc, *S. cerevisiae*; Mt, *M. thermoautotrophicum*; Bst, *B. stearothermophilus*; Tmt, *T. thermophilus*; Nc, *N. crassa*; Scm and Ncm, mitochondrial enzymes of Sc and Nc, respectively.

11-amino acid "signature sequence" that ends in the tetrapeptide HIGH (19, 46) and the KMSKS (47) element, both of which form part of the ATP binding site (48, 49). The signature sequence is part of a highly conserved loop between β_A and helix HA, whereas KMSKS is between two β-strands (Fig. 1).

The nucleotide-binding fold is divided by segments [designated connective polypeptides 1 and 2 (CP1 and CP2)] (22) that in the related class I glutaminyl-tRNA synthetase provide sequence elements for docking the acceptor end of bound tRNA into the proximity of the adenylate (49). Mutations in CP1 of methionyl-tRNA synthetase affect the tRNA interaction, possibly by altering contacts with the tRNA acceptor helix (50, 51). After the amino-terminal nucleotide-binding fold, there is a carboxyl-terminal helix-rich domain. The alignment of the 15 sequences in Fig. 1 supports the earlier conclusions (20, 22, 24) that major portions of isoleucyl-tRNA synthetase are folded like *E. coli* methionyl-tRNA synthetase.

The multiple sequence alignment shows that there are more frequent insertions into isoleucyl-tRNA synthetase than into methionyl-tRNA synthetase, accounting in part for the greater size of the former protein. Thus, among other considerations, isoleucyl-tRNA synthetase provided an opportunity to explore the possibility that these insertions occur between substructures that make up entire domains and, therefore, to determine whether reassembly is more likely to occur when breaks are introduced at these places. The coding sequence for *ileS* was divided at 18 places that included the apparent, idiosyncratic insertions (Fig. 2). The fragment pairs are designated P101/103, M157/157, etc., and give the last amino acid of the amino-terminal piece and the first amino acid of the carboxyl-terminal segment. The largest overlap among these pairs is the G330/329 construction (2 amino acids) and the largest deletion is contained in the S207/211 pair (3 amino acids deleted).

Functional Assembly of Split Proteins. The resulting amino-terminal and carboxyl-terminal portion of each pair was cloned behind a *lacZ* promoter and expressed separately from plasmids pKS148 (see *Materials and Methods*) and pTZ19R (42), respectively. These constructions give a carboxyl-terminal "tail" of 4–9 extraneous amino acids (encoded by pKS148) on the amino-terminal fragments, and an amino-terminal fusion of 15–29 amino acids (from the beginning of the β-galactosidase protein that is encoded by pTZ19R) on the carboxyl-terminal fragments. These tails are indicated in Fig. 2. Thus, complementation by individual fragments or by fragment combinations requires flexibility that is sufficient to displace the tails that are located at the junctions of the split sequences.

The split sites extend from codon 101 to codon 886 of the 939-amino acid polypeptide (Fig. 1). The activity of these fragments *in vivo* was individually tested by introduction of each plasmid into *E. coli* strain IQ843/pRMS711. This

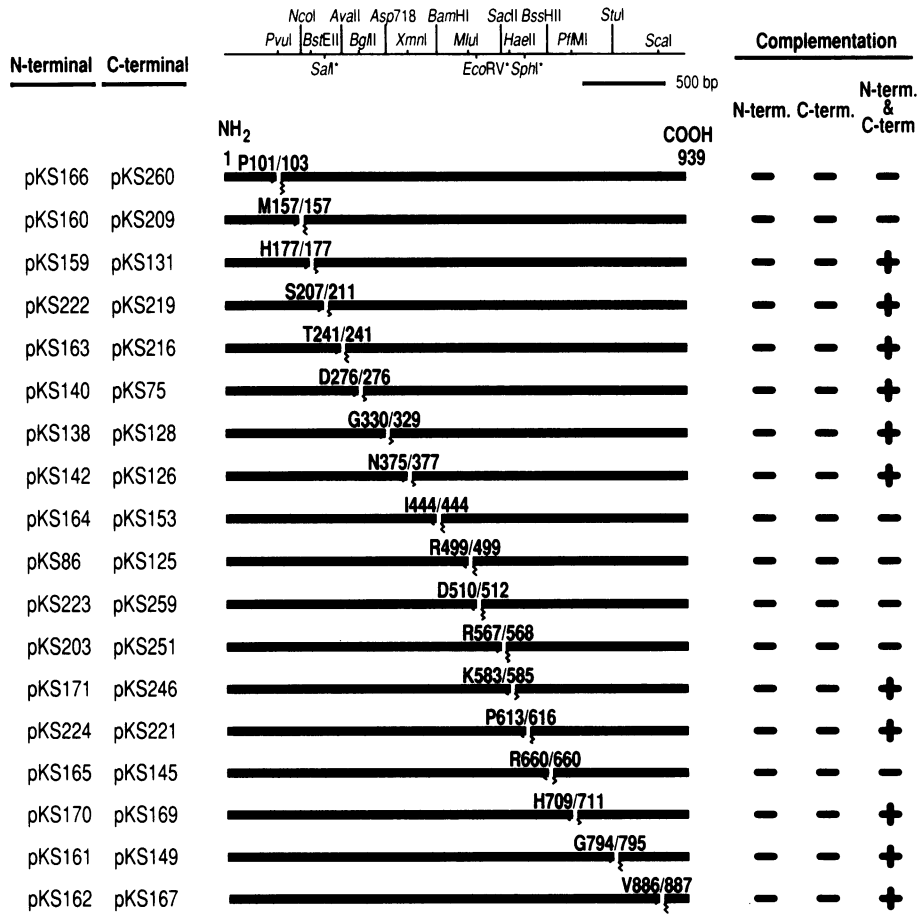

FIG. 2. Reconstitution of isoleucyl-tRNA synthetase activity by fragment pairs. The top line illustrates the *ileS* coding region and indicates the restriction sites that were used. Restriction sites that were introduced by site-directed mutagenesis are marked with an asterisk. Scale bar represents 500 base pairs (bp). DNA fragments producing amino-terminal and carboxyl-terminal fragments were cloned into plasmid vectors pKS148 and pTZ19R, respectively. Fragments produced are represented by thick lines and the residue numbers of the amino-terminal and carboxyl-terminal parts are given. Tails at the break points are also schematically shown, as zigzag lines. Plasmids were introduced into *E. coli* testor strain IQ843, which was then selected for tetracycline- and/or ampicillin-resistant transformants at 30°C. Transformants were scored for growth on M9 plates containing 0.2% Casamino acids or on LB plates at 42°C. Results are shown at right.

"testor" strain, constructed specifically for these studies by use of previously described methods (44, 52), has a complete ablation of *ileS*. Because *ileS* is an essential gene, cell viability is maintained by plasmid pRMS711, which encodes wild-type *ileS* and which has a temperature-sensitive replicon. As a result, growth of IQ843 is temperature-sensitive. Rescue of the temperature-sensitive phenotype can occur only when plasmids are introduced that provide isoleucyl-tRNA synthetase activity at the restrictive temperature. However, none of the plasmids that encode the amino- or carboxyl-terminal fragments complemented strain IQ843 (Fig. 2).

The amino- and carboxyl-terminal plasmid sets contain replicons that belong to different compatibility groups [p15A (pKS148) and ColE1 (pTZ19R)] and, therefore, can coexist in the same cell. Thus, by introduction of the appropriate plasmids into strain IQ843, the two protein fragments of each split pair can be synthesized in trans from separate mRNAs. When tested for assembly of isoleucyl-tRNA synthetase activity, 11 of the 18 split pairs were active [Fig. 1 (filled arrowheads) and Fig. 2]. Because of possible interference caused by the extraneous carboxyl- and amino-terminal tails that are located at the fragment junctions, no conclusion can be drawn for the 7 combinations [Fig. 1 (open arrowheads)] that yielded no complementation. Proteins from cells that expressed the 11 active split pairs were analyzed by SDS/polyacrylamide gel electrophoresis and by Western blot analysis with anti-isoleucyl-tRNA synthetase antibodies. Fragments of the expected sizes were detected in most cases, with no evidence for the presence of the full-length enzyme (Fig. 3*A*). Gel filtration under nondenaturing conditions of extracts of cells representing six examples of active split pairs (H177/177, D276/276, G330/329, N375/377, P613/616, and G794/795) showed that isoleucyl-tRNA synthetase activity migrated at the identical position as the intact, monomeric wild-type enzyme. The N375/377 and P613/616 examples are shown in Fig. 3*B*.

Interpretation in Terms of Structural Model. Reconstitution of isoleucyl-tRNA synthetase activity (filled arrowheads in Fig. 1) is usually observed when splits are made within insertions that cannot be aligned with some or all of the other sequences. These include splits at D276/276, N375/377, L709/711, G794/795, and V886/887. The division at V886/887 yields a carboxyl-terminal fragment of just 53 amino acids that has been suggested as encoding a zinc-binding motif in the enzyme (37, 53). These insertions may indicate locations of domains or subdomains that can spontaneously associate through sequence complementarity at the domain interfaces.

However, fragment complementation is not limited to interactions between the interfaces of presumed domains. There are four instances where reconstitution occurs when a split is made within a well-conserved sequence motif (H177/177, T241/241, G330/329, and P613/616), and two in a less conserved region (S207/211 and K583/585). One of these split points (G330/329) occurs in a sequence element that is ablated from the methionyl- and cysteinyl-tRNA synthetases and is known to be in a region of isoleucyl-tRNA synthetase that is dispensable for activity (22). In contrast, the splits at K583/585 and P613/616 occur directly in the nucleotide-binding fold. The active P613/616 split is located in an α-helix that is adjacent to the critical KMSKS motif that is conserved among all ten class I aminoacyl-tRNA synthetases. Juxtaposed tails of 9 and 19 amino acids at the breakpoint are apparently displaced by the strength of the interaction between flanking complementary sequences. The reconstituted size and *in vivo* activity of the P613/616 split enzyme are evident in Fig. 3*B*. We estimate that the specific activity of the P613/616 split enzyme is ≈10% that of the wild-type protein.

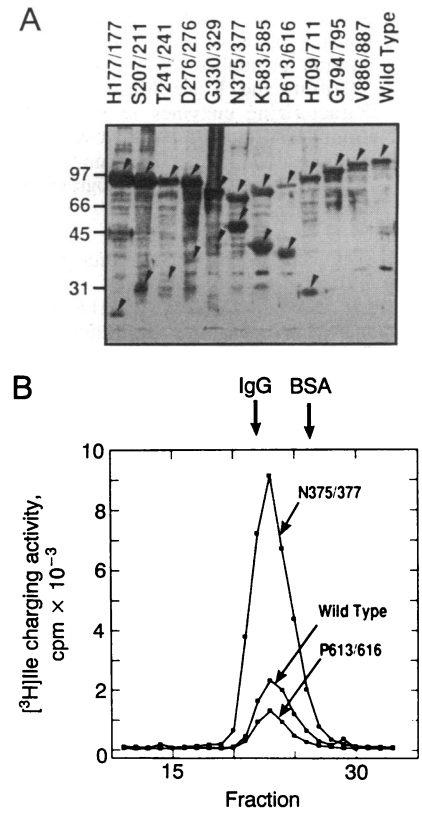

FIG. 3. (*A*) Western blot analysis of isoleucyl-tRNA synthetase fragments that were produced in complemented cells. Locations of size markers (Bio-Rad) are indicated at left ($M_r \times 10^{-3}$). Fragments of the expected sizes are indicated by filled arrowheads. The short fragments 795–939 and 887–939 could not be detected under these conditions. (*B*) Size fractionation of proteins under nondenaturing conditions. Strain IQ843 with the maintenance plasmid pRMS711 was used as the wild-type control. Because plasmid pRMS711 has a low copy number, the activity of wild-type protein is somewhat underrepresented relative to the N375/377 and P613/616 complementary pairs that are expressed from high-copy plasmids. The elution positions of size markers are shown at the top [IgG, M_r 148,000; bovine serum albumin (BSA), M_r 66,000].

Concluding Remarks. Neither overlapping fragments nor breaks at well-defined junctions between structural units/domains were required to achieve reconstitution in the present studies. The successful reconstitution of active monomeric enzyme from a large number (eleven) of essentially nonoverlapping fragment pairs, the presence of extraneous sequences at fragment junctions, and the ability to interrupt conserved regions of the structure suggest that internal complementarity is widely distributed throughout the entire sequence and is generally sufficient to overcome breaks in the peptide backbone to allow assembly of a functional structure. In addition to sequence-specific hydrogen bonding and electrostatic interactions, strong hydrophobic forces (54–58) that are broadly distributed and sequence-specific (59) probably contribute significantly to internal complementarity. Because separate recombinant plasmids were used to express fragment pairs of isoleucyl-tRNA synthetase in trans, internal complementarity must also be highly self-selective to overcome competing interactions with other protein sequences. These results potentially have implications for the evolution and assembly of large protein structures starting from specific noncovalent associations between individual polypeptide elements (60).

We thank Professors David Eisenberg (University of California, Los Angeles) and Peter Kim (Whitehead Institute and Massachusetts

Institute of Technology) and Dr. Alyssa Shepard (Massachusetts Institute of Technology) for reading and commenting on the manuscript before submission for publication. We thank Drs. Jonathan J. Burbaum, Linda Griffin, and Karin Musier-Forsyth for helpful suggestions in characterizing enzymes *in vitro*. This work was supported by National Institutes of Health Grant GM23562, by a grant from the Human Frontier Science Program, and by a fellowship to K.S. from the Human Frontier Science Program and the Toyobo Biotechnology Foundation.

1. Kato, I. & Anfinsen, C. B. (1969) *J. Biol. Chem.* **244**, 1004–1007.
2. Richards, F. M. & Wyckoff, H. W. (1971) *Enzymes* **4**, 647–806.
3. Anfinsen, C. B., Cuatrecasas, P. & Taniuchi, H. (1971) *Enzymes* **4**, 177–204.
4. Corradin, G. & Harbury, H. A. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 3036–3039.
5. Strittmatter, P., Barry, R. E. & Corcoran, D. (1972) *J. Biol. Chem.* **247**, 2768–2775.
6. Li, C. H. & Bewley, T. A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1476–1479.
7. Craik, C., Buchman, S. R. & Beychok, S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1384–1388.
8. Holmgren, A. & Slaby, I. (1979) *Biochemistry* **18**, 5591–5599.
9. Saint Girons, I., Gilles, A.-M., Margarita, D., Michelson, S., Monnot, M., Fermandjian, S., Danchin, A. & Bârzu, O. (1987) *J. Biol. Chem.* **262**, 622–629.
10. Galakatos, N. G. & Walsh, C. T. (1987) *Biochemistry* **26**, 8475–8480.
11. Bibi, E. & Kaback, H. R. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4325–4329.
12. Wrubel, W., Stochaj, U., Sonnewald, U., Theres, C. & Ehring, R. (1990) *J. Bacteriol.* **172**, 5374–5381.
13. Burbaum, J. J. & Schimmel, P. (1991) *Biochemistry* **30**, 319–324.
14. Taniuchi, H. & Anfinsen, C. B. (1971) *J. Biol. Chem.* **246**, 2291–2301.
15. Taniuchi, H., Parker, D. S. & Bohnert, J. L. (1977) *J. Biol. Chem.* **252**, 125–140.
16. Hantgan, R. R. & Taniuchi, H. (1977) *J. Biol. Chem.* **252**, 1367–1374.
17. Matsuyama, S., Kimura, E. & Mizushima, S. (1990) *J. Biol. Chem.* **265**, 8760–8765.
18. Hall, J. G. & Frieden, C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3060–3064.
19. Webster, T., Tsai, H., Kula, M., Mackie, G. A. & Schimmel, P. (1984) *Science* **226**, 1315–1317.
20. Hou, Y.-M., Shiba, K., Mottes, C. & Schimmel, P. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 976–980.
21. Brunie, S., Zelwer, C. & Risler, J. L. (1990) *J. Mol. Biol.* **216**, 411–424.
22. Starzyk, R. M., Webster, T. A. & Schimmel, P. (1987) *Science* **237**, 1614–1618.
23. Clarke, N. D., Lien, D. C. & Schimmel, P. (1988) *Science* **240**, 521–523.
24. Burbaum, J. J., Starzyk, R. M. & Schimmel, P. (1990) *Proteins* **7**, 99–111.
25. Needleman, S. B. & Wunsch, C. D. (1970) *J. Mol. Biol.* **48**, 443–453.
26. Feng, D.-F. & Doolittle, R. F. (1987) *J. Mol. Evol.* **25**, 351–360.
27. Gribskov, M., Lüthy, R. & Eisenberg, D. (1990) *Methods Enzymol.* **183**, 146–159.
28. Englisch, U., Englisch, S., Markmeyer, P., Schischkoff, J., Sternbach, H., Kratzin, H. & Cramer, F. (1987) *Biol. Chem. Hoppe-Seyler* **368**, 971–979.
29. Jenal, U., Rechsteiner, T., Tan, P.-Y., Bühlmann, E., Meile, L. & Leisinger, T. (1991) *J. Biol. Chem.* **226**, 10570–10577.
30. Härtlein, M., Frank, R. & Madern, D. (1987) *Nucleic Acids Res.* **15**, 9081–9082.
31. Jordana, X., Chatton, B., Paz-Weisshaar, M., Buhler, J.-M., Cramer, F., Ebel, J.-P. & Fasiolo, F. (1987) *J. Biol. Chem.* **262**, 7189–7194.
32. Borgford, T. J., Brand, N. J., Gray, T. E. & Fersht, A. R. (1987) *Biochemistry* **26**, 2480–2486.
33. Tzagoloff, A., Akai, A., Kurkulos, M. & Repetto, B. (1988) *J. Biol. Chem.* **263**, 850–856.
34. Chow, C. M., Metzenberg, R. L. & RajBhandary, U. L. (1989) *Mol. Cell. Biol.* **9**, 4631–4644.
35. Benarous, R., Chow, C. M. & RajBhandary, U. L. (1988) *Genetics* **119**, 805–814.
36. Dardel, F., Fayat, G. & Blanquet, S. (1984) *J. Bacteriol.* **160**, 1115–1122.
37. Nureki, O., Muramatsu, T., Suzuki, K., Kohda, D., Matsuzawa, H., Ohta, T., Miyazawa, T. & Yokoyama, S. (1991) *J. Biol. Chem.* **266**, 3268–3277.
38. Tzagoloff, A., Vambutas, A. & Akai, A. (1989) *Eur. J. Biochem.* **179**, 365–371.
39. Walter, P., Gangloff, J., Bonnet, J., Boulanger, Y., Ebel, J.-P. & Fasiolo, F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2437–2441.
40. Eriani, G., Dirheimer, G. & Gangloff, J. (1991) *Nucleic Acids Res.* **19**, 265–269.
41. Chang, A. C. Y. & Cohen, S. N. (1978) *J. Bacteriol.* **134**, 1141–1156.
42. Mead, D. A., Szczesna-Skorupa, E. & Kemper, B. (1986) *Protein Eng.* **1**, 67–74.
43. Silhavy, T. J., Berman, M. L. & Enquist, L. W. (1984) *Experiments with Gene Fusions* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), p. xi.
44. Toth, M. J. & Schimmel, P. (1986) *J. Biol. Chem.* **261**, 6643–6646.
45. Ghosh, G., Pelka, H. & Schulman, L. D. (1990) *Biochemistry* **29**, 2220–2225.
46. Ludmerer, S. W. & Schimmel, P. (1987) *J. Biol. Chem.* **262**, 10801–10806.
47. Hountondji, C., Blanquet, S. & Lederer, F. (1985) *Biochemistry* **24**, 1175–1180.
48. Mechulam, Y., Dardel, F., Le Corre, D., Blanquet, S. & Fayat, G. (1991) *J. Mol. Biol.* **217**, 465–475.
49. Rould, M. A., Perona, J. J., Söll, D. & Steitz, T. A. (1989) *Science* **246**, 1135–1142.
50. Meinnel, T., Mechulam, Y., Le Corre, D., Panvert, M., Blanquet, S. & Fayat, G. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 291–295.
51. Schimmel, P. (1991) *Curr. Opinion Struct. Biol.* **1**, 811–816.
52. Jasin, M. & Schimmel, P. (1984) *J. Bacteriol.* **159**, 783–786.
53. Miller, W. T., Hill, K. A. W. & Schimmel, P. (1991) *Biochemistry* **30**, 6870–6876.
54. Kauzman, W. (1959) *Adv. Protein Chem.* **14**, 1–63.
55. Chothia, C. (1976) *J. Mol. Biol.* **105**, 1–14.
56. Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H. & Zehfus, M. (1985) *Science* **229**, 834–838.
57. Nakai, K., Kidera, A. & Kanehisa, M. (1988) *Protein Eng.* **2**, 93–100.
58. Kellis, J. T., Jr., Nyberg, K., Săli, D. & Fersht, A. R. (1988) *Nature (London)* **333**, 784–786.
59. Shortle, D., Stites, W. E. & Meeker, A. K. (1990) *Biochemistry* **29**, 8033–8041.
60. Gilbert, W. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52**, 901–905.