There is much to learn about substrate recognition by ribozymes.

REFERENCES AND NOTES

1. R. B. Waring, C. Scazzocchio, T. A. Brown, R. W. Davies, *J. Mol. Biol.* **167**, 595 (1983); M. D. Been and T. R. Cech, *Cell* **47**, 207 (1986).
2. F. X. Sullivan and T. R. Cech, *Cell* **42**, 639 (1985); A. J. Zaug and T. R. Cech, *Science* **231**, 470 (1986).
3. A. J. Zaug, C. A. Grosshans, T. R. Cech, *Biochemistry* **27**, 8924 (1988).
4. N. Sugimoto, M. Tomka, R. Kierzek, P. C. Bevilacqua, D. H. Turner, *Nucleic Acids Res.* **17**, 355 (1989).
5. A. M. Pyle and T. R. Cech, *Nature* **350**, 628 (1991).
6. P. C. Bevilacqua and D. H. Turner, *Biochemistry* **30**, 10632 (1991).
7. R. Kierzek, P. C. Bevilacqua, Y. Li, D. H. Turner, in preparation.
8. The fluorescence intensities of unbound pyrCUCU, pyrCCUCU, and pyrCUCUCU are approximately 0.5, 0.65, and 0.5 that of pyrCCCUCU, respectively. Steady-state fluorescence measurements were performed on a Perkin-Elmer MPF-44A fluorimeter with excitation at 329 nm and emission at 397 nm under the conditions listed in Table 1.
9. S.-H. Kim and T. R. Cech, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8788 (1987); F. Michel and E. Westhof, *J. Mol. Biol.* **216**, 585 (1990).
10. Reactions were run as described (*6*) except in 5.0 mM $MgCl_2$, 135 mM NaCl, and 50 mM Hepes (25 mM Na^+) at pH 7.4, and reactions were initiated by addition of L-21 Sca I. Reactions went to approximately 90% completion. Formation of the product, p*UCG, was fit to [p*UCG] = [p*UCG]$_{t\,=\,\infty}$ $(1 - e^{-k_{obs}t})$, where k_{obs} is the observed rate constant (s^{-1}) and t is time (s). Because [UCGA] << [L-21 Sca I] << K_m^{UCGA} [K_m (Michaelis constant) for UCGA is 55 μM], k_{obs}/[L-21 Sca I]$_o$ = k_{cat}/K_m for UCGA. The k_{cat}/K_m for cleavage of 10 nM p*UCGA by pyrCUCU, CUCU, pyrCCUCU, CCUCU, and water (50 μM 5′ exon mimic and 5 μM L-21 Sca I) are 120, 120, 220, 130, and 55 M^{-1} s^{-1}, respectively. The k_{cat}/K_m for cleavage of 10 nM p*UCGA by pyrCCUCU, pyrCUCUCU, pyrCCCUCU, and water (0.5 μM 5′ exon mimic and 0.5 μM L-21 Sca I) are 140, 98, 68, and 58 M^{-1} s^{-1}, respectively. The presence of a ribose substrate suppresses the rate of hydrolysis in a related reaction (*4*).
11. K. Johnson, in *The Enzymes*, D. Sigman, Ed. (Academic Press, New York, 1992), vol. 20, pp. 1–61.
12. B. A. Barshop, R. F. Wrenn, C. Frieden, *Anal. Biochem.* **130**, 134 (1983); C. T. Zimmerle and C. Frieden, *Biochem. J.* **258**, 381 (1989).
13. Titration was performed with [L-21 Sca I] held constant at 40 nM; [pyrCUCU] was varied from 0.2 to 2.1 μM. Data were fit by nonlinear least squares to ΔF = [pyrCUCU]/($K_{-1}K_{-2}$ + [pyrCUCU]), where F is fluorescence intensity; $\Delta F \equiv F_{bound}/F_{bound}^{max}$, and $F_{bound} \equiv F_{total} - F_{oligomer}$. Here F_{bound}, F_{bound}^{max}, and $F_{oligomer}$ are, respectively, F due to bound oligomer at any [oligomer], F when all L-21 Sca I is bound to oligomer, and F due to unbound oligomer.
14. S. M. Freier *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9373 (1986); L. He, R. Kierzek, J. SantaLucia, Jr., A. E. Walter, D. H. Turner, *Biochemistry* **30**, 11124 (1991).
15. D. H. Turner, N. Sugimoto, S. M. Freier, in *Nucleic Acids*, vol. 1C of *Landolt-Bornstein Series*, W. Saenger, Ed. (Springer-Verlag, Berlin, 1990), pp. 201–227.
16. M. Eigen and G. G. Hammes, *Adv. Enzymol.* **25**, 1 (1963); C.-T. Lin, W. Böttcher, M. Chou, C. Creutz, N. Sutin, *J. Am. Chem. Soc.* **98**, 6536 (1976).
17. K. Yoon, D. H. Turner, I. Tinoco, Jr., *J. Mol. Biol.* **99**, 507 (1975); ———, F. von der Haar, F. Cramer, *Nucleic Acids Res.* **3**, 2233 (1976).
18. A. M. North, *The Collision Theory of Chemical Reactions in Liquids* (Methuen, London, 1964), chap. 7; K. Hiromi, *Kinetics of Fast Enzyme Reactions* (Kodansha, Tokyo, 1979), pp. 258–259; M. Ben-Nun and R. D. Levine, *J. Phys. Chem.* **96**, 1523 (1992).
19. D. Herschlag, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 6921 (1991).
20. The rate constant k_1 is thought to be limited by formation of a nucleation core of one to three nucleotides and should therefore be relatively independent of length beyond a trimer [J. Wetmur and N. Davidson, *J. Mol. Biol.* **31**, 349 (1968); D. Pörschke and M. Eigen, *ibid.* **62**, 361 (1971); M. E. Craig, D. M. Crothers, P. Doty, *ibid.*, p. 383; D. Pörschke, O. C. Uhlenbeck, F. H. Martin, *Biopolymers* **12**, 1313 (1973); (*15, 28*)]. Thus, increased binding is reflected primarily in a reduced k_{-1}.
21. Fitting of kinetic traces by simulation (*12*) suggests the intermediate for pyrCCUCU has only 60% of the fluorescence enhancement of the final state. Thus, the amplitude of the faster rate representing the intermediate is partially diminished relative to the total amplitude.
22. The rate plot for pyrCCCUCU is linear to at least 6.5 s^{-1}, somewhat greater than the expected asymptote for $1/\tau_2$, $k_2 + k_{-2}$, of 2.5 s^{-1}, on the basis of pyrCCUCU. An increase in k_2 from pyrCCUCU to pyrCCCUCU could cause this effect; this is unlikely, however, because tertiary interactions are confined to CUCU (*5, 6, 9*). Alternatively, because k_{-1} << (k_2, k_{-2}) for the hexamers (*14, 20*), if the intermediate and final species have similar fluorescence, Scheme I would be described by the simple rate $1/\tau = k_1$[pyrCCCUCU] + $k_{-1}k_{-2}/k_2$. This equation results in linear plots, as observed (Fig. 3B). The pyrCCCUCU has a large unbound fluorescence (*8*) and a small total change in fluorescence, suggesting that pyrCCCUCU may have different relative fluorescence coefficients for the intermediate and final species than pyrCCUCU or pyrCUCUCU.
23. D. Herschlag and T. R. Cech, *Biochemistry* **29**, 10159 (1990).
24. In this model, the faster transient described by Eq. 1 is not observed for pyrCUCU because the intermediate state does not build up because of the high value of K_{-1}.
25. Alternatively, binding of pyrCUCU to L-21 Sca I is also consistent with a one-step bimolecular mechanism with a slow association rate of 5.8 × 10^4 M^{-1} s^{-1} (*20*). Fluorescence-detected stopped-flow experiments with pyrCUCU and GGAGGA at 5°C (the signal-to-noise ratio was too low to permit a measurement at 15°C) indicate that the oligomers associate with a rate of 4.0 × 10^7 M^{-1} s^{-1} (Table 2). This is similar to the association rate of 5.5 × 10^7 M^{-1} s^{-1} for pyrCCUCU and GGAGGA at 15°C (Table 2), as expected because association rates for oligomers with GC pairs typically have low activation energies (*15, 28*). Thus, there is no strong dependence of k_1 on oligomer length for the model systems (*20*). The data for binding to L-21 Sca I, however, require a 70-fold increase in k_1 if the mechanism switches from one to two steps when pyrCUCU is replaced by pyrCCUCU (*29*). Because a large change in k_1 and a change in mechanism are unlikely, we favor Scheme I for pyrCUCU binding to L-21 Sca I.
26. R. C. Yuan, J. A. Steitz, P. B. Moore, D. M. Crothers, *Nucleic Acids Res.* **7**, 2399 (1979); D. Labuda, G. Striker, D. Pörschke, *J. Mol. Biol.* **174**, 587 (1984).
27. D. Herschlag, *Biochemistry* **31**, 1386 (1992). J1/2 is the unpaired A_3 stretch joining helix P1 and helix P2.
28. A. P. Williams, C. E. Longfellow, S. M. Freier, R. Kierzek, D. H. Turner, *ibid.* **28**, 4283 (1989).
29. Pulse-chase experiments with 5′ ^{32}P-labeled p*CUCU and L-21 Sca I at 50 mM Mg^{2+} and 50 mM Hepes (25 mM Na^+), pH 7.4, at 15°C give apparent association and dissociation rates similar to those for pyrCUCU in 5 mM Mg^{2+} (P. C. Bevilacqua and D. H. Turner, unpublished results). No trapping was observed at 5 mM Mg^{2+}. However, k_1 is expected weakly to be dependent on [Mg^{2+}] in this range (*28*).
30. K. Johnson, *Methods Enzymol.* **134**, 677 (1986).
31. M. Petersheim and D. H. Turner, *Biochemistry* **22**, 256 (1983).

32. We thank E. Kool, L. Lindahl, E. Phizicky, and K. Yoon for comments on the manuscript. Supported by the Office of Naval Research and NIH grants GM 22939 (to D.H.T.) and GM44613 (to K.A.J.). P.C.B. is a Messersmith Fellow and D.H.T. is a Guggenheim Fellow.

28 May 1992; accepted 1 September 1992

Three-Dimensional Structure of Dimeric Human Recombinant Macrophage Colony-Stimulating Factor

Jayvardhan Pandit, Andrew Bohm, Jarmila Jancarik, Robert Halenbeck, Kirston Koths, Sung-Hou Kim*

Macrophage colony-stimulating factor (M-CSF) triggers the development of cells of the monocyte-macrophage lineage and has a variety of stimulatory effects on mature cells of this class. The biologically active form of M-CSF is a disulfide-linked dimer that activates an intrinsic tyrosine kinase activity on the M-CSF receptor by inducing dimerization of the receptor molecules. The structure of a recombinant human M-CSF dimer, determined at 2.5 angstroms by x-ray crystallography, contains two bundles of four α helices laid end-to-end, with an interchain disulfide bond. Individual monomers of M-CSF show a close structural similarity to the cytokines granulocyte-macrophage colony-stimulating factor and human growth hormone. Both of these cytokines are monomeric in their active form, and their specific receptors lack intrinsic tyrosine kinase activity. The similarity of these structures suggests that the receptor binding determinants for all three cytokines may be similar.

Macrophage colony-stimulating factor (M-CSF) is one of a group of at least 18 glycoproteins, collectively known as hematopoietic growth factors, that regulate the growth and differentiation of blood cells (*1*). Native M-CSF is a 70- to 90-kD homodimer that stimulates proliferation and supports survival and differentiation of cells of the mononuclear phagocyte series (*2*). It also potentiates the ability of mature mononuclear phagocytes to perform their differentiated functions by enhancing their

Table 1. Diffraction data.

Parameter	Native	$K_2Hg(SCN)_4$	$Pt(NH_2Cl)_2$
Resolution (Å)	2.0	2.0	2.5
Total observations (no.)	65,871	62,645	34,985
Unique reflections (no.)	18,842	19,021	9,190
R_{merge} (%)	7.01*	6.9†	5.8†

*$I > 0$. †$I > 1\sigma_I$.

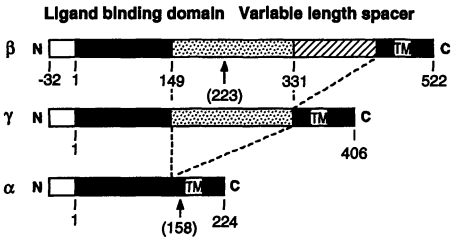


Fig. 1. Three biologically active forms of M-CSF. The open boxes (signal sequence), filled boxes, and transmembrane boxes (TM) are identical in all three forms. The stippled boxes represent sequences common to M-CSFβ and M-CSFγ, and the striped box represents coding sequence unique to M-CSFβ. Proteolytic cleavage sites, by which the hormone is presumed to be released from the membrane, are indicated by arrows.

ability to kill infectious microorganisms and tumor cells (*3*). In clinical trials, it has shown promise in the correction of blood cell deficiencies and fungal infections that arise as a side effect of chemotherapy and radiation therapy for cancer (*3*). In adult animals, M-CSF exerts its pleiotropic effects by binding a high-affinity cellular receptor. The M-CSF receptor is encoded by the *c-fms* proto-oncogene and is a member of a family of growth factor receptors that exhibit ligand-induced, tyrosine-specific protein kinase activity (*4*).

In contrast to other hematopoietic growth factors, M-CSF is biologically active only in a disulfide-linked dimeric form. By cDNA cloning, three forms of M-CSF have been found, M-CSFα (*5*) [256 amino acids (aa)], M-CSFβ (*6*) (554 aa), and M-CSFγ (*7*) (438 aa). All three forms are identical in their NH_2- and COOH-terminal regions, but M-CSFβ and M-CSFγ contain inserts at residue 149 as a result of alternative mRNA processing (Fig. 1). Fully active forms of M-CSF have been produced by using constructs that express as few as the first 150 aa of M-CSFα, demonstrating that the COOH-terminal region is not required for in vitro activity (*8*). The COOH-terminus does contain a transmembrane region that is believed to anchor some forms of M-CSF in an active, cell-associated form. The native secreted form of M-CSFα is generated by proteolytic cleavage of the membrane-bound form at or near residue 158. Here we report the backbone structure of an active, non-glycosylated and truncated sequence (residues 4 to 158) of M-CSFα determined at 2.5 Å resolution by x-ray crystallography.

Recombinant human M-CSF was purified from *Escherichia coli* and renatured in vitro to form a dimeric protein containing seven disulfide bonds (*9*). The specific biological activity of this non-glycosylated, truncated M-CSF is equal to that of glycosylated native M-CSF purified from human urine or recombinant M-CSFα secreted from CHO (Chinese hamster ovary) cells (*10*). The discovery that a dimeric M-CSF molecule containing seven to nine disulfide bonds could be properly refolded in vitro was not entirely expected.

A modified sparse matrix sampling method (*11*) was used to search for crystallization conditions. The best crystals were grown by vapor diffusion against 0.2 M $MgCl_2$, 100 mM tris-HCl, pH 8.5, and 23% polyethylene glycol 4000. The crystallization solution contained equal volumes of the reservoir and protein solution (10 mg/ml). The crystals have an orthorhombic space group $P2_12_12_1$, with unit cell parameters of a = 33.54, b = 65.26, and c = 159.63 Å, and diffract to 2.0 Å resolution when exposed to synchrotron radiation.

Intensity data were collected with imaging plates mounted on a Weissenberg camera modified for macromolecular crystallography (*12*) at the Photon Factory in Tsukuba, Japan. All data were collected with 1.00 Å wavelength x-rays and were processed with the program package WEIS (*13*). Two crystal settings were used to collect native data to a nominal resolution of 2.0 Å. Two derivatives were prepared by soaking crystals in heavy atoms dissolved in the reservoir solution. Both soaked crystals diffracted strongly, potassium mercury thiocyanate derivative to 2.0 Å and *trans*-dichloro-diammine platinum derivative to 2.5 Å. Isomorphous and anomalous difference Patterson maps clearly revealed one site for the Hg and two sites for the Pt derivative. Both derivatives proved to be non-isomorphous at high resolution and provided reliable phase information only to 3.0 Å. Anomalous and isomorphous phase information was used in initial phase refinement with the PROTEIN program package (*14*). The final figure of merit was 0.62 (10.0 to 3.0 Å, 6960 reflections, Tables 1 and 2). After solvent flattening (*15*), two bundles of four α helices related by an approximate twofold axis could be seen in the electron density map. Rotation and translation parameters of this non-crystallographic axis were refined by a density correlation method (*16*). Phases were then iteratively refined by molecular averaging and solvent flattening (*17*) with a manually defined mask for averaging (*18*). Chain tracing and model building were done in the resulting map [with the program FRODO (*19*)] keeping the original map as a reference.

The starting partial model for refinement contained only a polyalanine backbone for eight helices making up the two

Table 2. Phase refinement.

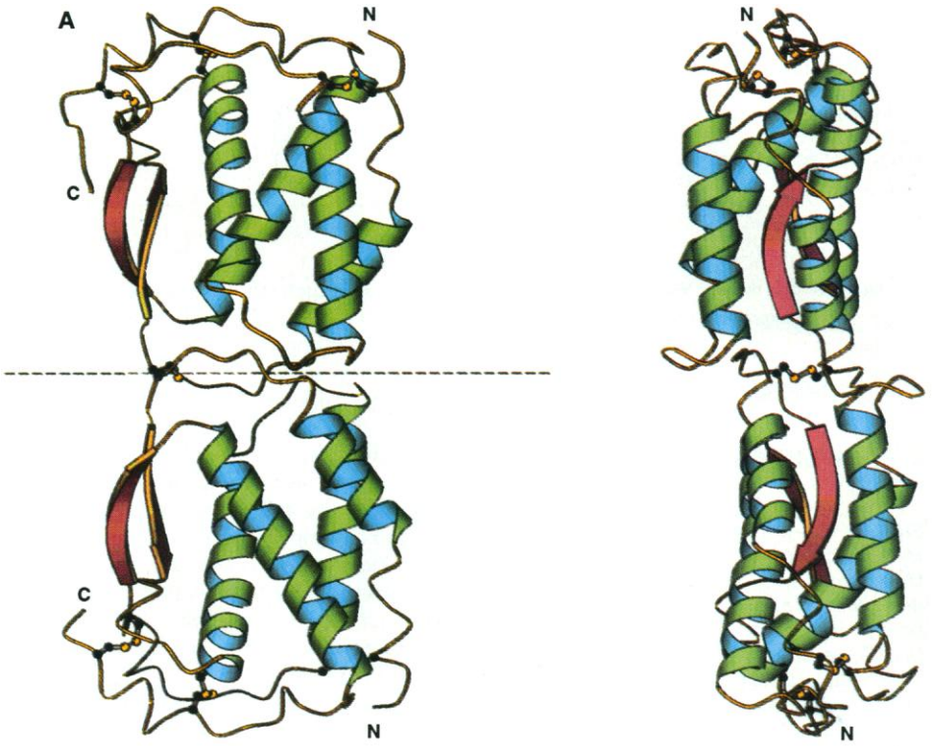
Parameter	Resolution (Å)								Overall
	7.7	6.3	5.3	4.6	4.1	3.6	3.3	3.0	
				Native					
Reflections (no.)	190	365	534	699	928	1149	1430	1665	6960
Figure of merit	0.58	0.67	0.66	0.64	0.64	0.61	0.59	0.62	0.62
				$K_2Hg(SCN)_4$					
R^*_{Cullis}	0.59	0.74	0.77	0.75	0.82	0.79	0.82	0.67	0.77
Phasing power†	1.54	1.46	1.22	1.15	0.93	0.86	0.97	1.50	1.16
				$Pt(NH_2Cl)_2$					
R_{Cullis}	0.76	0.69	0.59	0.74	0.80	0.76	0.73	0.93	0.77
Phasing power	1.13	1.70	1.87	1.16	0.99	1.03	1.16	0.95	1.33

*Cullis R factor for centric reflections. †Mean value of heavy atom structure factor amplitude divided by residual lack-of-closure error.

J. Pandit, Structural Biology Division, Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, CA 94720.
A. Bohm, Biophysics Graduate Group, University of California, Berkeley, CA 94720.
J. Jancarik, Department of Chemistry, University of California, Berkeley, CA 94720.
R. Halenbeck and K. Koths, Biological Therapeutics Department, Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608.
S.-H. Kim, Structural Biology Division, Lawrence Berkeley University of California, Berkeley, CA 94720.
*To whom correspondence should be addressed.

bundles. Positional refinement with the program XPLOR (*20*) gave an *R* factor of 0.49 to 3.0 Å. Partial model phase combination with the refined multiple isomorphous replacement phases resulted in a map of sufficient quality to allow the tracing of two long connected stretches of density running parallel to the four-helical bundle and a short loop connecting two of the helices. Two strong peaks in the density, one at the top of the first helix and the second lying directly on the molecular twofold axis, were assigned as disulfide-bonded cysteines. The number of residues between these two peaks uniquely identified the position in the sequence of these cysteines and consequently the sequence of the intervening residues. This initial registration was confirmed by the presence of a number of regions of strong density corresponding to aromatic side chains in the sequence. Partial model phase combination with the added loops and those side chains that were visible allowed the remaining residues to be registered, thus determining the overall topology of the molecule. The presence of seven disulfide bonds in the dimer, whose connectivity was unknown, made the initial chain tracing difficult but, once correctly assigned, provided a strong confirmation for the correctness of the tracing.

Further cycles of model building and refinement were carried out, eventually with the use of data from 8.0 to 2.5 Å resolution [$F > 2\sigma(F)$] to yield an *R* factor of 0.197. Forty-five water molecules, clearly visible in ($F_o - F_c$) maps, are part of the current model. The model at the present stage of refinement has root-mean-square (rms) deviations from ideal geometry of 0.019 Å for bond lengths and 3.7° for bond angles. The two monomers related by noncrystallographic symmetry have an rms deviation of 0.55 Å for 145 α-carbon atoms and 0.37 Å for 64 α-carbon atoms in the helical bundles alone. Eleven side chains are still not visible in the density and have been excluded from the current model. This model is also missing residues 154 to 158 in molecule A and residues 149 to 158 in molecule B at the COOH-terminus. The biological activity of M-CSF has been shown to be unaffected by the deletion of these ten residues (*10*), and it is interesting to note that residue 149 demarks an intron-exon border that may divide functional domains. It has been proposed (*7*) that M-CSF is expressed in variable lengths from a single gene in order to provide membrane-associated molecules with variable length tethers (Fig. 1). The membrane-associated form of M-CSF is also known to be biologically active and can functionally interact with receptor-bearing targets by direct cell-cell contact (*21*). The missing COOH-terminal residues probably represent the



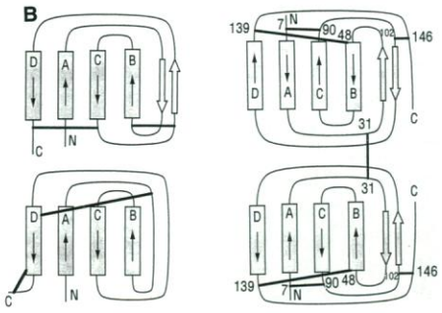
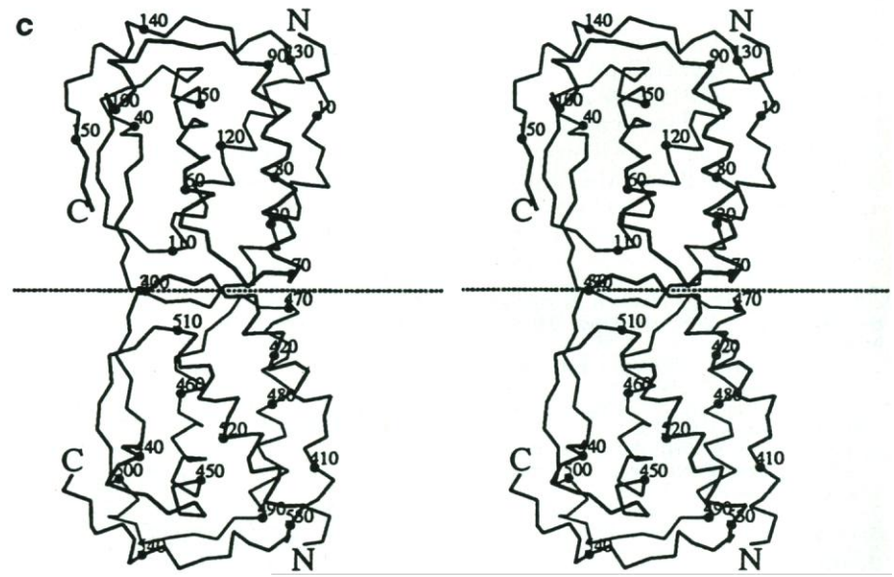


Fig. 2. Backbone structure of M-CSF. (**A**) Two views of a ribbon diagram highlighting the secondary structural elements: helices are shown as spiral ribbons and β strands as arrows (*31*). All of the Cys residues have been represented by a ball-and-stick model. N and C mark the termini. (Left) The noncrystallographic twofold axis is indicated by a dashed line. (Right) A view down the molecular twofold axis. (**B**) Topological structure, showing all the disulfide bonds. N and C mark the termini. Topological structures of GM-CSF (upper left) and GH (lower left) are shown for comparison. (**C**) A stereo Cα diagram with every tenth residue labeled. The noncrystallographic axis is shown as a dotted line.



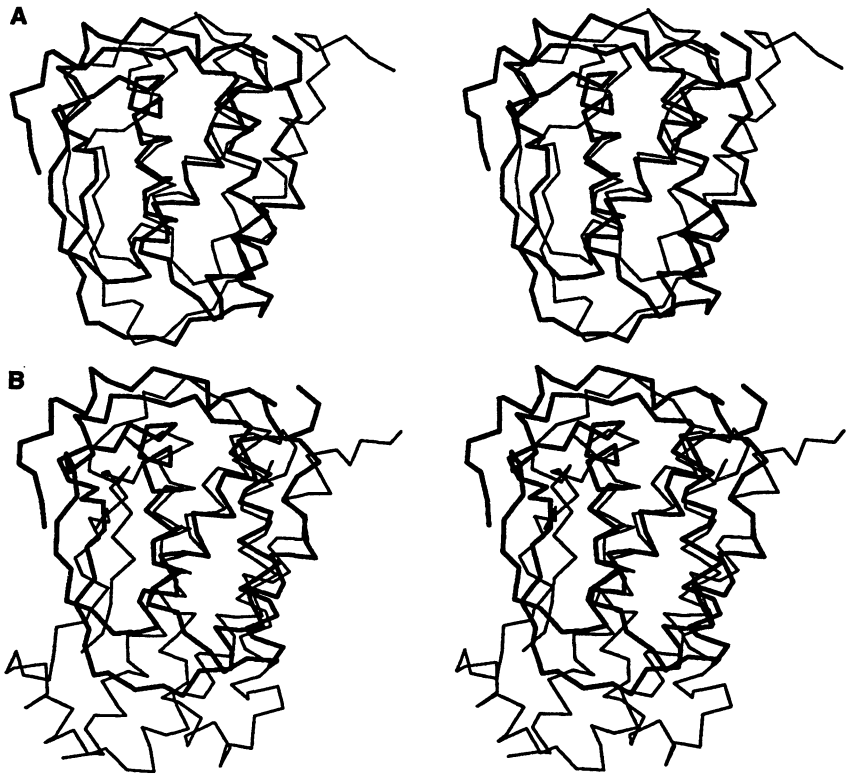


Fig. 3. Stereo views showing the superposition of Cα atoms of M-CSF (residues 4 to 149, thick lines) with (**A**) GM-CSF (residues 5 to 123, thin lines) and (**B**) GH (residues 1 to 190, thin lines). See Table 3.

Table 3. The residues used for calculating the best superposition and the root-mean-square (rms) deviations in their Cα positions. The total rms deviation for 51 residues between M-CSF and GM-CSF is 1.8 Å, and for 66 residues between M-CSF and GH is 1.7 Å.

Secondary structure	M-CSF/GM-CSF	M-CSF/GH
Helix A	13-22/20-29	13-26/11-24
Helix B	53-61/58-66	52-63/74-85
Helix C	80-89/73-82	71-90/108-127
Helix D	111-123/104-116	110-129/166-185
β-Strand I	35-39/40-44	
β-Strand II	102-105/97-100	

beginning of the flexible tether region.

M-CSF is known to have a covalently linked dimer as the active receptor-binding moiety. Our analysis reveals that the dimer is formed by linking two four-helix bundles end-to-end, forming an extremely flat, elongated structure (approximate dimensions 80 by 30 by 20 Å, Fig. 2, A and C). There are three intramolecular disulfide bonds in each monomer (Cys^{7}-Cys^{90}, Cys^{48}-Cys^{139}, and Cys^{102}-Cys^{146}), all of which are at the ends distal to the dimer interface. One interchain disulfide bond (Cys^{31}-Cys^{31}) is located at the dimer interface with the noncrystallographic twofold symmetry axis passing through it. The total surface buried by dimerization is ~850 $Å^2$ from each monomer, which constitutes about 9% of the total exposed area of a monomer (*22*).

Mutation experiments (*8*) have shown that the first six Cys residues in each chain are essential for biological activity. Our structure suggests that these cysteines are important for structure formation or stability rather than being directly involved in receptor recognition.

The overall topology of an M-CSF monomer is that of an antiparallel four-α-helical bundle, in which the helices run up-up-down-down (Fig. 2B). This structure is unlike the more commonly observed up-down-up-down connectivity of most four-helical bundles (*23*). A long crossover connection (residues 25 to 45) links helix A to helix B and a similar connection (residues 88 to 109) is found between helices C and D. Residues 33 to 38 from the first connecting strand and 103 to 108 from the second strand form an antiparallel β ribbon that fits into one "side" of the four-helix bundle. It is interesting that all other examples of four-α-helical bundle proteins that have this connectivity are also cytokines: granulocyte-macrophage colony-stimulating factor (GM-CSF) (*24*), human growth hormone (GH) (*25*), and porcine growth hormone (PGH) (*26*). The solution structure of another cytokine, interleukin-4, as determined by nuclear magnetic resonance spectroscopy, also shows the same topology (*27*).

Although M-CSF has no significant sequence homology with GM-CSF or GH, comparison with the recently published structures of these cytokines shows a remarkably conserved pattern of secondary and tertiary structure (Fig. 3 and Table 3). Superposition of M-CSF and GM-CSF reveals that in addition to the four α helices, the short β ribbon can also be overlapped in the two structures. Differences in the three structures are in the relative lengths of the helices and the connecting loops. Also, the Cys cross-links in all three structures are completely different (Fig. 2B), despite the overall topology being identical. In all three proteins, the four-helical bundle is coded for by a gene made up of four exons. The exon-intron junctions occur at nearly the same positions in the three-dimensional structure in all three cases, at the end of helix A, at the beginning of helix B, at the end of helix C, and at the end of helix D.

Bazan (*28*) aligned the sequences of M-CSF and stem-cell factor (SCF) by registering equivalent exons in the respective genes and then matching exon-encoded secondary structure by predictive algorithms. His prediction of a GH-like topology for M-CSF is borne out by the present study. Interestingly, SCF lacks three of the seven cysteines found in M-CSF. Two of these, Cys^{102} and Cys^{146}, are disulfide-linked to each other in M-CSF, and the third, Cys^{31}, makes an intermolecular disulfide bridge between M-CSF monomers. It would seem that the end-to-end packing of the four-helix bundles seen in M-CSF may, in SCF, form a stable dimer, even in the absence of the intermolecular disulfide bond due to the relatively large contact surface area (~850 $Å^2$) between the monomers.

The extracellular domains of a number of cytokine receptors are structurally and perhaps evolutionarily related (*29*). This hematopoietic growth factor receptor family does not have a tyrosine kinase activity, and includes receptors for GM-CSF; G-CSF; interleukins-2, -3, -4, -5, -6, and -7; GH; prolactin; and erythropoietin. The M-CSF receptor, on the other hand, belongs to another family, the subclass III receptor tyrosine kinases. These receptors bind dimeric forms of M-CSF, SCF, and platelet-derived growth factor (PDGF), leading to kinase activation and phosphorylation of many substrates involved in the production of second messengers (*30*).

The fact that M-CSF adopts the same folding motif as the other cytokines, and indeed has a close structural similarity, suggests that analogous regions on the hor-

mone may be involved in receptor binding for both classes of receptors. By analogy with GH and GM-CSF, this would indicate that the exposed surface of helix D or the residues on the "side" of the bundle on helices A and C are most likely to be implicated in receptor binding.

It seems that a four-helical bundle with the distinctive "double-overhand" topology is a common structural scaffolding adopted by all the cytokines that bind receptors belonging to the hematopoietic superfamily and subclass III receptor tyrosine kinases. This structural motif is preserved across a diverse array of hormones that often possess no detectable sequence homology.

REFERENCES AND NOTES

1. N. A. Nicola, *Annu. Rev. Biochem.* **58**, 45 (1989); S. C. Clark and R. Kamen, *Science* **236**, 1229 (1987).
2. E. R. Stanley, L. J. Guilbert, R. J. Tushinsky, S. H. Bartelmez, *J. Cell. Biochem.* **21**, 151 (1983).
3. P. Ralph *et al.*, *Immunobiology* **172**, 194 (1986).
4. C. J. Sherr *et al.*, *Cell* **41**, 665 (1985).
5. E. S. Kawasaki *et al.*, *Science* **230**, 291 (1985).
6. G. G. Wong *et al.*, *ibid.* **235**, 1504 (1987).
7. D. P. Cerretti *et al.*, *Mol. Immunol.* **25**, 761 (1988).
8. E. S. Kawasaki and M. B. Ladner, *Immunol. Ser.* **49**, 155 (1990).
9. R. Halenbeck, E. Kawasaki, J. Wrin, K. Koths, *BioTechnology* **7**, 710 (1989).
10. R. Halenbeck and K. Koths, unpublished observations.
11. J. Jancarik and S.-H. Kim, *J. Appl. Crystallogr.* **24**, 409 (1991).
12. N. Sakabe, *ibid.* **16**, 542 (1983).
13. T. Higashi, *ibid.* **22**, 9 (1989).
14. W. Steigemann, thesis, Technische Universität München (1974).
15. B. C. Wang, *Methods Enzymol.* **115**, 90 (1985).
16. J. M. Cox, *J. Mol. Biol.* **28**, 151 (1967). The relation between the two molecules is described by the spherical angles ϕ = 25.2°, ψ = 167.2°, κ = 180.0°, and a translation of 141.5, −5.9, and 361.2 Å, respectively, in an orthogonal coordinate system described by *a*, *b*, and *c*. The noncrystallographic symmetry axis was refined with program RHOREF written by P. Chakravarti and D. Rees at UCLA.
17. G. Bricogné, *Acta Crystallogr.* **32**, 832 (1976).
18. An envelope for noncrystallographic symmetry averaging was determined by inspecting the electron density map on the graphics screen and filling up the regions of density believed to correspond to one dimer with dummy atoms.
19. T. A. Jones, *J. Appl. Crystallogr.* **11**, 268 (1978).
20. A. T. Brünger, *XPLOR, A System for Crystallography and NMR, 2.1 Manual* (Yale University, New Haven, CT, 1990).
21. J. Stein, G. V. Borzillo, C. W. Rettenmeier, *Blood* **76**, 1308 (1990).
22. B. Lee and F. M. Richards, *J. Mol. Biol.* **55**, 379 (1971). Solvent accessibility was calculated by using Lee and Richards' algorithm with a probe radius of 1.4 Å.
23. S. R. Presnell and F. E. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6592 (1989).
24. K. Diederichs, T. Boone, P. A. Karplus, *Science* **254**, 1779 (1991).
25. A. M. de Vos, M. Ultsch, A. A. Kossiakoff, *ibid.* **255**, 306 (1992).
26. S. S. Abdel-Meguid *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6434 (1987).
27. R. Powers *et al.*, *Science* **256**, 1673 (1992).
28. J. F. Bazan, *Cell* **65**, 9 (1991).
29. ———, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6934 (1990).
30. Y. Yarden and A. Ullrich, *Annu. Rev. Biochem.* **57**, 443 (1988).
31. Figures 2, A and C, and 3 were prepared with the program MOLSCRIPT [P. J. Kraulis, *J. Appl. Crystallogr.* **24** (1991)].
32. We thank P. A. Karplus and A. M. deVos for GM-CSF and GH coordinates, respectively. We gratefully acknowledge Chiron Corporation, National Science Foundation (BBS 87-20137), and Department of Energy for a grant, equipment support, and computing support, respectively. The Cα coordinates will be deposited in the Brookhaven Protein Data Bank.

15 June 1992; accepted 20 August 1992

Consequences of Replication Fork Movement Through Transcription Units in Vivo

Sarah French

To examine the basis for the evolutionary selection for codirectionality of replication and transcription in *Escherichia coli*, electron microscopy was used to visualize replication from an inducible ColE1 replication origin inserted into the *Escherichia coli* chromosome upstream (5′) or downstream (3′) of *rrnB*, a ribosomal RNA operon. Active *rrnB* operons were replicated either in the same direction in which they were transcribed or in the opposite direction. In either direction, RNA polymerases were dislodged during replication. When replication and transcription were codirectional, the rate of replication fork movement was similar to that observed in nontranscribed regions. When replication and transcription occurred in opposite directions, replication fork movement was reduced.

RNA polymerases (RNAPs) and DNA polymerases (DNAPs) are highly processive enzyme complexes that use the same DNA template to guide their activities. Unless transcription and replication are spatially or temporally separated, there may be times in the cell cycle when RNAPs and DNAPs must compete for right-of-way along the same stretch of DNA. Replication in *Escherichia coli* is initiated bidirectionally from a single site, *oriC*. In rich medium, the time it takes to replicate the chromosome is longer than the doubling time of the cell (*1*). In order for each daughter cell to receive a full chromosomal complement under these conditions, a new round of replication must be initiated before the previous round is completed. Six or more replication forks may, therefore, be present on the chromosome at the same time. Rapid cell growth also requires high amounts of gene expression to provide adequate amounts of cellular precursors. Because DNA replication in *E. coli* proceeds at 20 times the rate of transcription (*2*, *3*), collisions between RNAPs and DNAPs can occur. How are these conflicts resolved?

The bacterial chromosome apparently is organized to avoid head-on collisions between DNAPs and RNAPs (*4*, *5*). Codirectionality of replication and transcription predominates, especially for genes that are frequently transcribed (*5*). Of the genes known to encode protein synthesis machinery, 90% are transcribed in the same direction in which they are replicated. It has been postulated that, if replication and transcription are codirectional, replication forks can follow slowly behind RNAPs until transcription is terminated (*5*). At that point, replication can continue at its normal speed. Head-on collisions would be more difficult to resolve, consistent with an apparent evolutionary selection against their occurrence.

Researchers' efforts to visualize the interaction between replication and transcription (*6*) have been subject to uncertainties about where the replication originated, what the transcriptional state of the DNA was before replication, and whether observed replication forks were moving or stationary. I have inserted here an inducible replication origin into the *E. coli* chromosome in a region that is well characterized, easily recognized, and frequently transcribed. Using the Miller chromatin spreading technique (*7*, *8*) in conjunction with electron microscopy, I observed the positions of replication forks relative to transcribing RNAPs after induction of replication.

The inducible replication origin was inserted upstream or downstream of *rrnB*. One of seven ribosomal RNA (rRNA) operons in *E. coli*, *rrnB* is a good template over which to view the interaction of replication and transcription because it is long [5.4 kb (*9*)] and actively transcribed [12 RNAPs per kilobase (*7*)]. The rRNA operons are easily recognized in chromatin spreads by their unique morphology (Figs. 1A and 2A). They are densely packed with RNAPs, and the processing of transcripts between the 16*S* and 23*S* cistrons (*10*) gives rise to two gradients of increasing transcript length. The *rrnB* operon can be identified by its proximity to the *tufB* transcription unit 4 kb downstream (*7*). No other *rrn* operon

Department of Biology, University of Virginia, Charlottesville, VA 22903.