Primary Structure of Neocarzinostatin, an Antitumor Protein

Abstract. *The antitumor protein neocarzinostatin is an acidic single-chain molecule, cross-linked by two disulfide bridges, and consists of 109 amino acid residues. Complete disulfide bond reduction and S-carboxymethylation was achieved in liquid ammonia. Sequence determination of five tryptic fragments led to the proposed primary structure.*

Most antibiotics of known structure are small to medium-sized molecules of molecular weights between a few hundred and a few thousand daltons (*1*). Several macromolecular antibiotics have attracted interest because of their antitumor activity and relatively low toxicity. These agents are small proteins that have been isolated from the culture filtrates of various microorganisms belonging to the genus *Streptomyces* (*2*). One if these is neocarzinostatin (NCS). It was isolated (*3*) from *Streptomyces carzinostaticus* var. F-41 and chemically characterized (*4*) as an acidic single-chain polypeptide. This compound exhibits typical antibiotic activity (*3*) against Gram-positive bacteria, such as *Sarcina lutea* [minimum inhibitory concentration (MIC), 2 μg/ml] or *Staphylococcus aureus* (MIC, 8 to 16 μg/ml), and is highly effective against several experimental tumors in mice, including ascitic sarcoma 180 (100 percent survival at daily doses of 0.8 to 3.2 mg/kg × 6, intraperitoneally), ascitic leukemia SN-36 (*3*), and leukemia L-1210 (*5*). The principal effect of NCS is arrest of mitosis (*6*), and the main molecular target in susceptive bacteria appears to be DNA. Studies on the mode of action (*7*) indicate that NCS inhibits DNA synthesis and also causes degradation of existing DNA (*7, 8*) in *Sarcina lutea*. Clinical tests are promising for the treatment of tumors in the rectum and stomach (*9*) and of carcinoma of the bladder and penis (*10*).

We now report the amino acid sequence of NCS molecule (*11*). It was purified from a commercial product (*12*) by twice-repeated carboxymethyl (CM)-cellulose chromatography through gradient or stepwise elution with 0.1*M* sodium acetate buffers at *p*H 3.2 and 3.5, followed by dialysis and Sephadex G-50 gel filtration in 0.1*M* acetic acid (*13*). The protein contains 109 amino acid residues, the molecular weight being approximately 10,700. The amino acid composition (*14*) is unusual in its high content of alanine, glycine, serine, and threonine (Table 1). NCS has four half-cystine and two tryptophan residues (*15*), but no histidine or methionine. Carboxypeptidase A digests (*16*) showed the carboxyl terminal to be -Ile-Ser-Phe-Asn (*17*). The amino terminal sequence, Ala-Ala-Pro-Thr-Ala-Thr-Val-Thr-, was determined by the Edman-dansyl procedure (*18, 19*).

Native NCS resists digestion by trypsin. Performic acid oxidation followed by incubation with trypsin gave heterogeneous peptides that were not useful for sequence analysis (*13, 20*). Many attempts at reductive disulfide bond cleavage failed. Neocarzinostatin is extraordinarily resistant to reduction by sodium borohydride, mercaptoethanol, or dithiothreitol even in the presence of the highest possible concentrations of the denaturing agents urea or guanidinium chloride (*13*). After the conventional procedures (*21*) proved to be inadequate for NCS, we found that treatment with sodium in liquid ammonia (*22*) effected rapid and quantitative reduction of the disulfide bond. However, undesired peptide bond cleavage also occurred, probably at proline residues via the known Birch reduction (*23*). We discovered subsequently (*24*) that complete disulfide reduction can readily be obtained by dithiothreitol (*25*) in liquid ammonia, without detectable side reactions. Alkyl chlorides effect full thiol alkylation in this solvent (*21, 26*). Evaporation and simple repeated washing of the residual solids with methanol directly afford homogeneous reduced and alkylated protein derivatives (*23*). For sequence analysis the *S*-carboxymethylated derivative of NCS was prepared by this procedure with chloroacetic acid as thiol alkylating agent. All four half-cystine residues were fully alkylated.

Tryptic digestion of *S*-carboxymethylated NCS at *p*H 8.2 showed five fragments on peptide mapping (*27*).

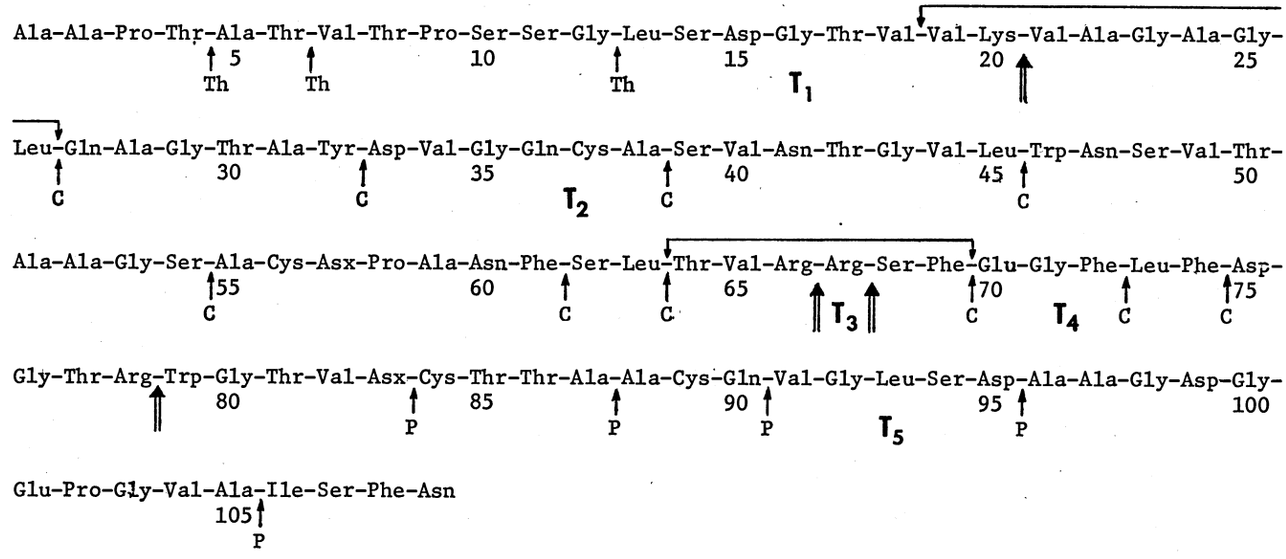

Fig. 1. Structure of neocarzinostatin. Double arrows indicate tryptic cleavage; small arrows show some of the subsequent cleavages by thermolysin (*Th*), chymotrypsin (*C*), or pepsin (*P*). Overlap peptides are indicated by a bridged arrow.

Table 1. Amino acid compositions and amino terminals of tetra-*S*-carboxymethyl neocarzinostatin and of the peptides isolated after digestion with trypsin. Amino acid analyses (*14*) were done on 20- and 48-hour hydrolyzates (6*N* HCl, 110°C) in a Phoenix model M-6800 analyzer.

Residue	T_1	T_2	T_3	T_4	T_5	CM-NCS
Lys	1.0 (1)					0.9 (1)
Arg		1.0 (1)	1.0 (1)	1.0 (1)		3.1 (3)
CM-Cys*		1.5 (2)			1.4 (2)	3.9 (4)
Asp	1.0 (1)	5.5 (5)		1.0 (1)	3.8 (4)	11.5 (11)
Thr	3.7 (4)	4.3 (4)		1.0 (1)	2.8 (3)	11.9 (12)
Ser	2.7 (3)	4.1 (4)		0.9 (1)	1.8 (2)	9.6 (10)
Glu		2.0 (2)		1.0 (1)	2.0 (2)	5.0 (5)
Pro	2.1 (2)	1.3 (1)			1.0 (1)	4.3 (4)
Gly	2.0 (2)	6.2 (6)		2.1 (2)	5.1 (5)	15.4 (15)
Ala	3.0 (3)	9.7 (9)			5.0 (5)	17.4 (17)
Val	2.8 (3)	5.8 (6)			3.1 (3)	12.5 (12)
Ile					1.1 (1)	1.1 (1)
Leu	0.9 (1)	3.0 (3)		1.0 (1)	1.1 (1)	6.0 (6)
Tyr		0.9 (1)				1.0 (1)
Phe		0.9 (1)		2.9 (3)	0.9 (1)	5.1 (5)
Trp		1.0 (1)†			1.0 (1)†	2.0 (2)‡
Total	20	46	1	11	31	109
Amino terminal§	Ala	Val	Arg	Ser	Trp‖	Ala

* *S*-Carboxymethylcysteine. † Determined spectrophotometrically. ‡ Determined colorimetrically after chemical modification (*15*). § Determined by dansylation (*19*). ‖ By reaction with fluorescein isothiocyanate.

Preparative separation and purification of the tryptic peptides was carried out by combinations of paper chromatography, paper electrophoresis, and Sephadex gel filtration. Amino acid compositions of the tryptic peptides are given in Table 1. Fragment T_3 is free arginine. The other fragments were further digested by thermolysine (T_1), by chymotrypsin (T_2 and T_4), and by pepsin (T_5). The separated subfragments were sequenced by standard procedures, in particular by the subtractive and Edman-dansyl procedures (Fig. 1). The amino terminal 25 residues of T_2 were also analyzed with an automated peptide sequenator (*28*). The arrangement of tryptic fragments, in the order shown, was based on (i) coincidence of amino and carboxyl terminal sequences of NCS with those of T_1 and T_5, respectively, and (ii) on identification of overlap peptides from a peptic-chymotryptic digest of tetra-*S*-carboxymethyl NCS. Isolation by paper chromatography and paper electrophoresis was followed by amino acid and sequence analysis. Overlap peptide Val-Lys-Val-Ala-Gly-Ala-Gly-Leu, positions 19 to 26, connected T_1 and T_2; and Thr-Val-Arg-Arg-Ser-Phe, positions 64 to 69, connected T_2 and T_4 with the single arginine T_3 placed in between at position 67.

In the proposed primary structure of NCS (Fig. 1), the single tyrosine residue is located in position 32, the tryptophan residues in positions 46 and 79; and several clusters of Ala, Gly, Ser, and Thr are conspicuous. The positions of the disulfide bonds remain to be determined (*29, 30*).

The knowledge of the NCS structure offers possibilities of chemical synthesis—for example, preparation of parts of the molecule in search for an active core (*31*). Group-specific chemical modification of the protein should be of particular interest. Deamination of NCS at the amino terminal or acylation by treatment with fluorescein isothiocyanate has been reported (*32*) to result in considerably decreased toxicity but retention of high antitumor activity in mice.

JOHANNES MEIENHOFER
HIROSHI MAEDA*, CHARLES B. GLASER
JOZSEF CZOMBOS, KENJI KUROMIZU
Children's Cancer Research Foundation and Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02115

References and Notes

1. See D. Gottlieb and P. D. Shaw, Eds., *Antibiotics* (Springer-Verlag, New York, 1967), vols. 1 and 2.
2. Protein antibiotics with antitumor activity include "carzinostatin" from *Streptomyces carzinostaticus* [N. Ishida, K. Kumagai, K. Miyazaki, M. Ito, *Gann* **51** (Suppl.), 56 (1960)]; "mitomalcin" from *S. malayensis* [T. C. McBride, M. Axelrod, W. P. Cullen, W. S. Marsh, C. S. Sodano, K. V. Rao, *Proc. Soc. Exp. Biol. Med.* **130**, 1188 (1969)]; "macromomycin" from *S. macromomyceticus* [H. Chimura, M. Ishizuka, M. Hamada, S. Hori, K. Kimura, J. Iwanaga, T. Takeuchi, H. Umezawa, *J. Antibiot.* **21**, 44 (1968)]; "actinoxanthin" from *S. globisporus* [N. M. Vichrova, T. I. Kruchkova, E. V. Preobrazhenskaya, A. S. Khokhlor, *Antibiotiki* **2**, 21 (1957)]; and others.
3. N. Ishida, K. Miyazaki, K. Kumagai, M. Rikimaru, *J. Antibiot.* **18**, 68 (1965).
4. H. Maeda, K. Kumagai, N. Ishida, *ibid.* **19**, 253 (1966).
5. W. T. Bradner and D. J. Hutchison, *Cancer Chemother. Rep.* **50**, 79 (1966).
6. K. Kumagai, Y. Ono, T. Nishikawa, N. Ishida, *J. Antibiot.* **19**, 69 (1966).
7. Y. Ono, Y. Watanabe, N. Ishida, *Biochim. Biophys. Acta* **119**, 46 (1966).
8. Y. Ono, Y. Ito, H. Maeda, N. Ishida, *ibid.* **155**, 616 (1968).
9. M. Takahashi, K. Toriyama, H. Maeda, M. Kikuchi, K. Kumagai, N. Ishida, *Tohoku J. Exp. Med.* **98**, 273 (1969).
10. K. Matsumoto and K. Nakauchi, in preparation.
11. A short report was presented [J. Meienhofer, H. Maeda, C. B. Glaser, K. Kuromizu, *Fed. Proc.* **31**, 445 (1972)]. Positions 39, 101/102, and some amide positions of this tentative structure require confirmation. The structure of polypeptide A from *S. carzinostaticus* bears no resemblance to NCS. It consists of 87 amino acid residues and has one disulfide bridge [H. Sato, T. Tanimura, T. Nakajima, Z. Tamura, *Chem. Pharm. Bull.* **17**, 2188 (1969)].
12. Neocarzinostatin was purchased from Kayaku Antibiotics Research Laboratory, Tokyo, Japan.
13. J. Meienhofer, H. Maeda, C. B. Glaser, J. Czombos, in *Progress in Peptide Research*, S. Lande, Ed. (Gordon & Breach, New York, 1972), pp. 295–306.
14. D. H. Spackman, W. H. Stein, S. Moore, *Anal. Chem.* **30**, 1190 (1958); D. H. Spackman, *Methods Enzymol.* **11**, 3 (1967).
15. H. Maeda and J. Meienhofer, *Int. J. Protein Res.* **2**, 135 (1970).
16. ———, *Fed. Eur. Biochem. Soc. Lett.* **9**, 301 (1970). The carboxy terminal sequence proposed at that time [Ser-Val-Ala-Ile-Phe-Asn] has been revised by transposing the Ser residue by three positions to the right.
17. Abbreviations, Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartic acid; Cys, cyteine; Glu, glutamic acid; Gln, glutamine; Gly, glycine; Ile, isoleucine; Leu, leucine; Lys, lysine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; Val, valine.
18. P. Edman, *Acta Chem. Scand.* **4**, 283 (1950).
19. W. R. Gray and B. S. Hartley, *Biochem. J.* **89**, 50P (1963); K. R. Woods and K.-T. Wang, *Biochim. Biophys. Acta* **133**, 369 (1967).
20. Performic acid–oxidized NCS was digested to give five main fragments, but efforts to obtain the tryptophan-containing peptides in homogeneous form failed. The heterogeneity was probably due to tryptophan oxidation during performic acid treatment. Therefore, this approach had to be abandoned [see (*13*)].
21. J. L. Bailey, *Techniques in Protein Chemistry* (Elsevier, New York, 1967), pp. 115–124; T. A. Bewley and C. H. Li, *Int. J. Protein Res.* **1**, 117 (1969).
22. V. du Vigneaud, L. F. Audrieth, H. S. Loring, *J. Amer. Chem. Soc.* **52**, 4500 (1930).
23. K. Hofmann and H. Yajima, *ibid.* **83**, 2289 (1961); M. Wilchek, S. Sarid, A. Patchornik, *Biochim. Biophys. Acta* **104**, 616 (1965); W. F. Benisek, M. A. Raftery, R. D. Cole, *Biochemistry*, **6**, 3780 (1967).
24. J. Meienhofer, J. Czombos, H. Maeda, *J. Amer. Chem. Soc.* **93**, 3080 (1971). The limitation of the procedure resides with solubility of a given peptide or protein in liquid ammonia; although many proteins are readily dissolved, others are insoluble in this medium.
25. W. W. Cleland, *Biochemistry* **3**, 480 (1964).
26. M. D. Armstrong and J. D. Lewis, *J. Org. Chem.* **16**, 749 (1951).
27. Chromatography: *n*-butanol–acetic acid–water (4 : 1 : 5, by volume); electrophoresis: formic acid–acetic acid–water, *p*H 1.9 (25 : 87 : 888, by volume); paper: Whatman 3 MM.
28. C. B. Glaser, H. Maeda, J. Meienhofer, H. Niall, *Experientia* **28**, 772 (1972).
29. C. B. Glaser, H. Maeda, J. Meienhofer, *J. Chromatogr.* **50**, 151 (1970).
30. H. Maeda, C. B. Glaser, J. Meienhofer, *Biochem. Biophys. Res. Commun.* **39**, 1211 (1970).
31. J. Meienhofer, J. Czombos, K. Kuromizu, in preparation.
32. K. Kumagai, H. Maeda, N. Ishida, *Antimicrob. Ag. Chemother.* **1966**, 546 (1967).
33. Supported in part by NCI research grant C-6516 and NIH grant FR-05526. We thank G. Johnson, J.. Shaw, J. Daniel, C. Moriarty, L. Gill, and E. Judkins for technical assistance.
* Present address: Department of Microbiology, Kumamoto University Medical School, Kumamoto, Japan.

16 June 1972