Vol. 190, No. 1, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1993 Pages 20-26

THE DISULFIDE BOND ARRANGEMENT OF LEUKEMIA INHIBITORY FACTOR: HOMOLOGY TO ONCOSTATIN M AND STRUCTURAL IMPLICATIONS

Nicos A. Nicola†, Bronwyn Cross† and Richard J.Simpson*

†The Walter and Eliza Hall Institute of Medical Research and Cooperative Research Centre for Cellular Growth Factors, P.O. Royal Melbourne Hospital, Parkville, Victoria 3050, Australia

*Joint Protein Structure Laboratory of The Walter and Eliza Hall Institute and the Ludwig Institute for Cancer Research, Melbourne Tumour Biology Branch, P.O. Royal Melbourne Hospital, Parkville, Victoria 3050, Australia

Received October 29, 1992

Murine leukemia inhibitory factor (LIF) (the fully active recombinant form produced in E.coli) was digested in the unreduced state with trypsin and Staphylococcal V8 protease in 0.05% sodium dodecyl sulfate. Disulfide-bonded peptides were identified by altered mobility on reverse-phase high-performance liquid chromatography in the presence or absence of dithiothreitol and subjected to amino acid sequencing. Peptides containing more than one disulfide bond were subjected to further proteolysis and disulfide-bonded subfragments identified and sequenced. The three disulfide bonds are CYS13-CYS135, CYS19-CYS132 and CYS61-164 and the first and third of these are clearly homologous to the two disulfide bonds in oncostatin M. The spatial organization of the cysteine residues in the predicted four α-helical bundle structure of LIF (Bazan, Neuron 7,197;1991) is compatible with these disulfide assignments.

Leukemia inhibitory factor (LIF) is a polyfunctional cytokine which inhibits the growth of some leukemic cell lines by inducing their terminal differentiation, inhibits the differentiation of totipotent embryonic stem cells, stimulates osteoblast function, inhibits lipoprotein lipase activity of adipocytes, co-stimulates megakaryocyte and platelet formation, induces the release of acute phase proteins from liver and has a variety of effects on neuronal cells (1-3). Some of these activities are shared by other cytokines including interleukin-6 (IL-6), granulocyte colony stimulating factor (G-CSF), oncostatin M (OSM) and ciliary neurotrophic factor (CNTF) (4,5). Indeed these cytokines appear to have weak homology in their amino acid sequences and a common predicted four α-helical bundle structure (4,6). Moreover, the receptors for each of these cytokines consist of one or more subunits with homologous extracellular domains (hemopoietin receptor domain) and several of these cytokines (LIF, IL-6, OSM, CNTF) share common receptor subunits involved in high-affinity binding and signal transduction (5,7,8).

Amongst these cytokines, LIF is most closely related to OSM (27% sequence identity) and the receptors for these cytokines share at least two common subunits (gp130) and the LIF receptor α-chain) (7,8). It has been suggested that OSM may possess all of the biological activities exhibited

0006-291X/93 $4.00

by LIF. Since there should be common structural elements in LIF and OSM involved in binding to the common receptor subunits and activation of the receptor it is important to determine the structure of LIF and compare it to OSM. Here we determine the disposition of the three disulfide bonds of LIF, show that two of these are homologous to the two disulfides in OSM, and demonstrate that the disulfide arrangement is consistent with the four α-helical bundle structure predicted for LIF and OSM.

MATERIALS AND METHODS

Recombinant murine LIF was produced in Escherichia coli as a fusion protein with glutathione-S-transferase separated by a thrombin cleavage site as described (9). Thrombin cleavage generated the N-terminal amino acid sequence Gly-Ser-Pro.... which has an extra one or two amino acids compared to native murine LIF which starts either at Ser or at Pro (10,11). Recombinant LIF was purified as described except that an additional chromatographic step on S-Sepharose (Pharmacia, Uppsala) was added between the affinity purification and reverse-phase high performance liquid chromatography (RP-HPLC) steps. The final product gave a single protein band of Mr 22,000 (>95% purity) on SDS PAGE and had a specific biological activity ($1\text{-}2\times10^8$ U/mg) indistinguishable from native murine LIF (12). Staphylococcal V8 protease (550 U/mg) was from ICN Biochemicals (Costa Mesa)) and TPCK-treated bovine pancreatic trypsin (10,000-13,000 BAEE Units/mg) was from Sigma (St. Louis). Dithiothreitol (DTT) was electrophoresis grade from BioRad (Richmond, Calif.), Acetonitrile and trifluoroacetic acid (TFA) were sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis HPLC grade from Mallinckrodt (Paris, Ky) and Pierce (Rockford, Ill.), respectively. Water was of MilliQ quality (Millipore, Milford, Mass).

In a first digestion, 500 µg of LIF in 200 µl sodium phosphate buffered (20 mM, pH 7.4) saline (0.15 M) (PBS) was mixed with 200 µl of 0.1 M Tris-HCl pH 8.0/0.1% sodium dodecyl sulfate (SDS) and 25 µg of TPCK-trypsin (12 µl of a solution in water) and 25 µg of V8 protease (25 µl of a solution in water) added. The solution was incubated at 37°C for 6 days then a further 25 µg each of TPCK-trypsin and V8 protease added and incubated at 37°C for a further 48 h. An aliquot (120 µl) of the reaction mixture was loaded onto a Brownlee RP-300 C8 column (250x7 mm, 7µm particle size) and eluted for 10 min with water/0.1% TFA (buffer A) at a flow rate of 1 ml/min followed by a 60 min linear gradient to 60% acetonitrile/0.085% TFA (buffer B) at the same flow rate on a Hewlett-Packard 1090 HPLC system. Individual peaks were collected by hand after allowing for the delay time between detector and outlet tube. A separate aliquot of the reaction mixture (20 µl) was incubated with 10 mM DTT for 5 min at 95°C and eluted under identical conditions to identify disulfide-linked peptide peaks. Two such peaks (A and B) in the original chromatogram were identified, diluted with an equal volume of water and further purified by a second chromatography on the same matrix but using a smaller column (30 x 2.1mm). The column was eluted with Buffer A for 5 min at 1 ml/min, then the flow rate reduced to 0.1 ml/min over 1 min and the column eluted with a 60 min linear gradient to 100% buffer B at 0.1 ml/min. Peaks were hand collected and subjected to amino acid sequencing.

In a second digestion 800 µg of a separate batch of LIF in 880 µl PBS was mixed with 880 µl of 0.1 M Tris-HCl buffer, pH 8/0.1% SDS and incubated with 40 µg each of TPCK-trypsin and V8 protease for 4 days at 37°C. A further 40 µg each of the two enzymes were added and incubation continued for 48 h at 37°C. 500 µl of this reaction mixture was separated as above except that the linear gradient was replaced by a 20 min linear gradient to 50% buffer B followed by a 60 min linear gradient to 100% buffer B. Peaks C and D were identified as disulfide-containing peptides and repurified as described above. 20% of the collected volume was used for amino acid sequencing.

The remaining 80% of peak C (160 µl) was evaporated to near dryness in a Speedi-Vac vacuum concentrator and 200 µl of 0.1 M $NaHCO_3$, pH 8 added. 10 µg of TPCK-trypsin (5 µl in the same buffer) was added and the mixture incubated at 37°C for 24 hr. Aliquots of this reaction mixture (20 µl each) with or without 10 mM DTT were chromatographed on the 30x2.1 mm C8 column as described above to identify disulfide-bonded peptides. Then 150 µl of reaction mixture was

chromatographed under identical conditions and DTT-sensitive peaks (E,F) collected by hand. These peaks were repurified on the same column (after dilution with an equal volume of water) and the purified peptides subjected to amino acid sequencing.

Amino acid sequencing was performed using an Applied Biosystems Model 470A protein sequenator equipped with a model 120A on-line phenylthiohydantoin amino acid analyser as previously described (13).

RESULTS AND DISCUSSION

Murine LIF in the folded state was resistant to proteolysis by trypsin and V8 protease in the absence of denaturants as assessed by the generation of new peaks on RP-HPLC or new peptide bands on Tricine-SDS gels (14). Even in the presence of 4M urea or 0.1% SDS little digestion of LIF was observed in the presence of trypsin (1:10 enzyme to substrate ratio). However significant digestion did occur when LIF was treated with V8 protease at pH7.0 in 0.01-0.1% SDS at an enzyme to substrate ratio of 1:10 and additional digestion bands were observed when trypsin was included in the reaction mixture. Apparently the V8 protease cleavages exposed cryptic tryptic cleavage sites.

A time course of digestion of murine LIF (500 μg) with V8 protease and trypsin in 100mM Tris-HCl pH8, 0.1% SDS indicated that a near limit digestion pattern (assessed by RP-HPLC) could be obtained by incubation at 37°C for 6 days (enzyme to substrate ratio 1:20 for each) followed by an increase of the enzyme to substrate ratio to 1:10 for each and incubation at 37°C for a further 48 h.

Fig. 1a shows the resulting digestion pattern of 500 μg LIF when applied to a RP-HPLC column (RP300 C8) and eluted with an acetonitrile gradient in 0.1% TFA (1% per min at 1 ml/min). All peaks were collected by hand and the peaks labeled A and B were identified as peptide peaks that altered their mobility when the digest was rechromatographed under identical conditions in the presence of 10 mM DTT (data not shown). Each peak was rechromatographed on a smaller RP300 C8 column (30x2.1 mm) in the absence of DTT utilizing the same acetonitrile/TFA gradient at 0.1 ml/min and peptides A and B collected as shown in Fig. 1b.

Fig. 2 shows that peak B contained three independent amino acid sequences linked by disulfides while peptide A contained two independent amino acid sequences. This latter peptide revealed an unequivocal disulfide bond between residues CYS61-CYS164.

Fig. 3 shows an identical analysis of murine LIF except that a larger amount (800 μg) of LIF was used in an attempt to obtain sufficient amounts of peptide for further digestion. The peaks labeled C and D were identified as peptides that altered their mobility when rechromatographed in the presence of DTT. Peptides C and D were rechromatographed as above for peptides A and C and subjected to amino acid sequencing (Fig.2). Both C and D contained three independent sequences that could be recognized as tryptic cleaves of murine LIF. Each contained the N-terminal peptide and the adjacent tryptic peptide beginning at position 17 (C2). Peak C contained in addition a

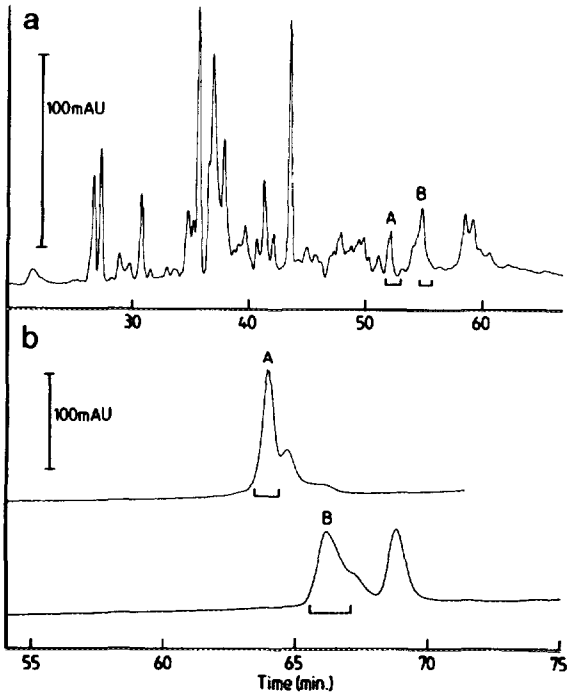


Figure 1. Chromatograph of the 500 µg digestion of LIF with TPCK-trypsin and V8 protease on a RP300 C8 HPLC column (250x7 mm). DTT-sensitive peaks (A,B) are shown in panel a. Peaks A and B were further purified on a smaller RP300 C8 column (30x2.1 mm) as shown in panel b. Pooling of peaks is indicated by the horizontal bars. The vertical bars represent milliabsorbance units (mAU) at 214 nm. The second major peak next to peak B was a peptide derived from V8 protease (ILPNNDRHQ).

tryptic peptide beginning at position 116 (C3) while peak D contained an overlapping tryptic peptide beginning at position 104 (D3). Since these peptides must be connected by disulfide bonds the latter peptides must extend past the two cysteine residues at positions 132 and 135 and disulfides must exist between C1 and C3 (D3) and between C2 and C3 (D3). However, it is not possible from these data to distinguish between the two disulfide possibilities C13-C132: C19-C135 and C13-C135: C19-C132.

In order to distinguish these possibilities it was necessary to cleave between CYS132 and CYS135. One of the two amino acids between these two cysteine residues was an arginine and we reasoned that trypsin might be able to cleave at this site if a higher enzyme to substrate ratio was used with pure peptide C in the absence of SDS. Peptide C was evaporated to near dryness and resuspended in 200 µl of 0.1M $NaHCO_3$ buffer at pH8. Ten µg of trypsin in the same buffer was added and incubated for 24 h at 37°C. The resulting digestion pattern is shown in Fig. 4 and two peaks (E,F) were identified where elution position was sensitive to DTT. Peak E yielded two amino acid sequences corresponding to the N-terminal sequence (C1) and the expected peptide LXNK - resulting from tryptic cleavage at arginine 133. This indicates a disulfide link between CYS13 and CYS135 (X in the above sequence). Peak F yielded the expected complementary

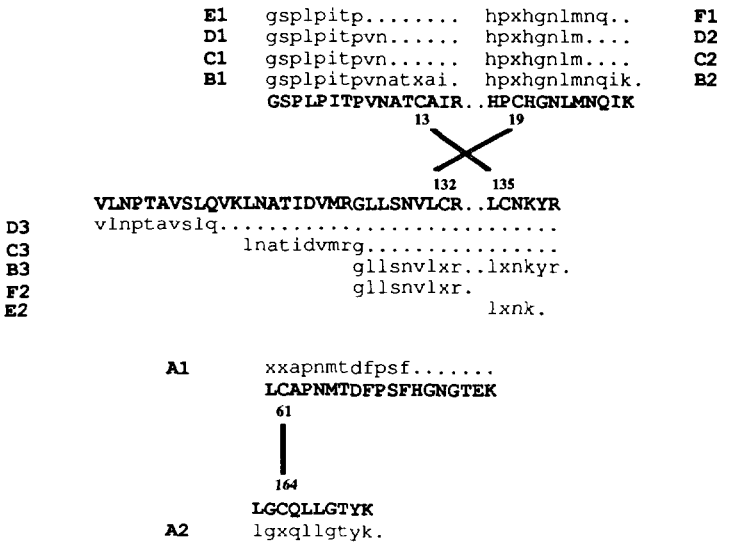


Figure 2. Amino acid sequences of DTT-sensitive peptide peaks. The amino acid sequence of murine LIF with final disulfide assignments is shown in the centre of each part of the diagram. The deduced sequences contained within peaks A,B,C,D,E and F are shown above or below this sequence in lower case. Peaks B,C,D contained three sequences each (e.g.,B1, B2, B3) while Peaks A,E and F contained two sequences each (e.g.,A1, A2). The gaps in the sequence at positions 17/18 and 133/134 are artificial to distinguish between peptides that span these regions or terminate at those sites. Amino acid residues are designated by the one-letter notation and X indicates an unidentified phenylthiohydantoin amino acid derivative in the sequence analysis.

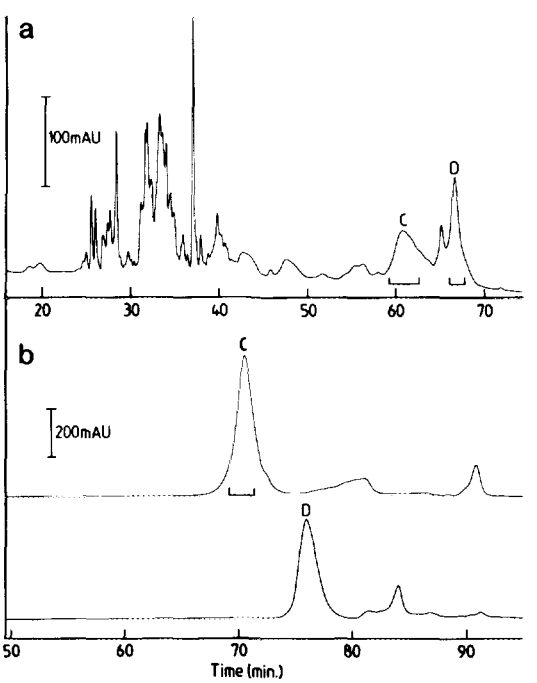
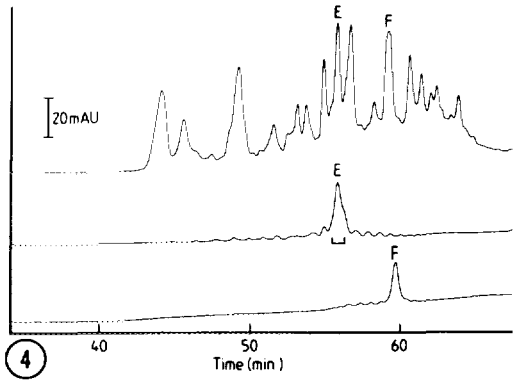


Figure 3. Chromatograph of the 800 µg digestion of LIF with TPCK-trypsin and V8 protease on a Brownlee RP300 C8 HPLC column (250x7 mm). In this case the gradient was 20 min to 30% CH_3CN/TFA followed by 60 min to 60% CH_3CN/TFA. DTT-sensitive peaks (C,D) are shown in panel a and were further purified on a 30x2.1 mm column as shown in panel b. The vertical bars represent mAU at 214 nm.



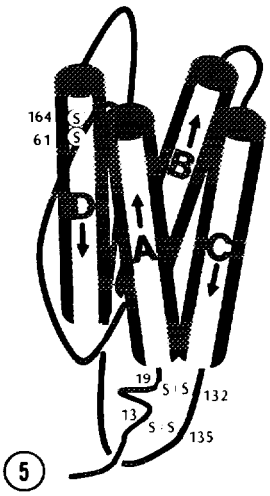


Figure 4. Peak C from Figure 3 was further digested with TPCK-trypsin in the absence of SDS and chromatographed on the smaller RP300 C8 column (30x2.1 mm). DTT-sensitive peaks (E,F) were repurified on the same column and subjected to amino acid sequencing. The vertical bars represent mAU at 214 nm.

Figure 5. Disposition of the three disulfide bonds based on the four α-helical bundle three-dimensional structure of LIF proposed by Bazan (6).

peptide sequences corresponding to C2 and the truncated peptide C3 beginning after arginine 124 (limit digest). This indicates a disulfide link between CYS19 and CYS132.

The disulfide arrangements deduced above are all internally consistent and no peptides with alternate disulfide arrangements were found. This suggests a unique folded conformation with no detectable misfolded molecules consistent with the wild type specific biological activity of the recombinant molecule. The disulfide arrangement is consistent with that proposed for OSM which displays 27% sequence identity with murine LIF. Alignment of the amino acid sequences of LIF and OSM suggest homologous pairing of disulfides: CYS6-126 (OSM) with CYS 13-135 (LIF) and CYS48-166 (OSM) with CYS61-164 (LIF) (OSM contains only two disulfide bonds and is lacking a homologous disulfide to LIF CYS 19-132).

These disulfide assignments for LIF are compatible with the predicted structure of LIF which consists of a four α-helical bundle with connecting loops (6) (Fig. 5). In this model the N-terminal tail containing CYS13 and CYS19 is juxtaposed in an antiparallel fashion to the loop between helices C and D which contains CYS132 and CYS135. Similarly CYS61 in the loop between helices A and B is closely aligned to CYS164 at the start of helix D. OSM has a similar predicted structure (4) and mutational analysis has suggested that only the disulfide bond corresponding to CYS61-CYS164 in LIF is essential for biological activity (15).

It has been suggested that LIF, OSM, CNTF, IL-6 and G-CSF may share some structural features in common (4,6). Indeed the receptors for each of these growth factors are related and belong to a

superfamily of cytokine receptors. Moreover, LIF, CNTF, OSM and IL-6 each share a common receptor subunit (gp130) in their receptor complexes (5,7,8). Since gp130 can confer high-affinity binding to each of these ligand-receptor complexes, it is likely that common structural features in each of the ligands is involved. It is hoped that further structural elucidation of these growth factors will help to define these common structural elements.

ACKNOWLEDGMENTS

This work was supported by the National Health and Medical Research Council, Canberra, Australia; the Australian Government Cooperative Research Centres Scheme; the National Institutes of Health, NCI Grant CA22556, Bethesda, MD; the Anti-Cancer Council of Victoria and the J.D. and L. Harris and Philip Bushell Trusts. We gratefully acknowledge the lead-up work performed on this project by Marcia Ward and Denis Quilici.

REFERENCES

1. Metcalf D. (1991) Int J Cell Cloning 9, 95-105.
2. Hilton DJ. (1992) Trends Biochem Sci 17, 72-76.
3. Kurzvock R, Estrov Z, Wetzler M, Gutterman JU and Talpaz M. (1991) Endocr Rev 12, 208-217.
4. Rose TM and Bruce AG. (1991) Proc Natl Acad Sci (USA) 88, 8641-8645.
5. Ip NY, Nye SH, Boulton TG, Davis S, Taga T, Li Y, Birren SJ, Yasukawa K, Kishimoto T, Anderson DJ, Stahl N and Yancopoulos GD. (1992) Cell 69, 1121-1132.
6. Bazan JF. (1991) Neuron 7, 197-208.
7. Gearing DP, Comeau MR, Friend DJ, Gimpel SD, Thut CJ, McGourty J, Brasher KK, King JA, Gillis S and Mosley B. (1992) Science 255, 1434-1437.
8. Gearing DP and Bruce AG. (1992) New Biol 4, 61-65.
9. Gearing DP, Nicola NA, Metcalf D, Foote S, Willson TA, Gough NM and Williams RL. (1989) Biotechnology 7, 1157-1161.
10. Simpson RJ, Hilton DJ, Nice EC, Rubira MR, Metcalf D, Gearing DP, Gough NM and Nicola NA. (1988) Eur J Biochem 175, 541-547.
11. Yamamori T, Fukada K, Aebersold R, Korsching S, Fann M-J and Patterson PH. (1989) Science 246, 1412-1416.
12. Hilton DJ, Nicola NA and Metcalf D. (1988) Anal Biochem 173, 359-367.
13. Simpson RJ, Moritz RL, Rubira MR, Gorman JJ and Van Snick J (1989) Eur J Biochem 183, 715-722.
14. Schagger H and von Jagow G. (1987) Anal Biochem 166, 368-379.
15. Kallestad JC, Shoyab M and Linsley PS. (1991) J Biol Chem 266, 8940-8945.