Vol. 95, No. 2, 1980 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 31, 1980 Pages 872-879

CHOU-FASMAN ANALYSIS OF THE SECONDARY STRUCTURE OF F AND LE INTERFERONS

Teresa G. Hayes

Department of Microbiology
New York University School of Medicine
New York, NY 10016

Received June 23,1980

SUMMARY: The Chou-Fasman method for calculating the secondary structure of F and Le interferons from their amino acid sequences predicts several regions of homologous structure in the two interferons. Those areas tend to be located in the middle and C- terminal ends of the molecules. Total predicted content of α helix is 55% for Le interferon and 36% for F interferon. Predicted β-pleated sheet residues total 33% for F and 16% for Le.

INTRODUCTION: Interferons are inducible proteins synthesized by mammalian cells in response to stimulation by viruses, synthetic polyribonucleotides, mitogens, and a variety of other substances. Their biological actions are diverse. In vitro, interferons have been shown to prevent virus multiplication, inhibit cell division, stimulate the activity of "natural killer" cytotoxic lymphocytes, and enhance phagocytosis by macrophages (1,2). They have shown promising clinical potential as future antiviral and/or antitumor agents (3).

In vivo trials have been limited by the necessity to use very high doses and by a shortage of purified material for administration. Production of large quantities of highly purified interferons is physically difficult and exceedingly expensive. Only recently was it possible to obtain enough of the purified proteins for partial N-terminal amino acid sequence analysis (4,5). Consequently, no studies have been done on the physical conformation of interferon molecules.

Of the three known types of human interferons, F interferon, the predominant product of poly(I)·poly(C)-induced fibroblast cells in tissue culture, has many properties in common with Le interferon, the main species made

0006-291X/80/140872-08$01.00/0

by virus-stimulated lymphoid cells. The third type, immune (or T) interferon, has been less well characterized and will not be analyzed in this study. F and Le interferons are glycoproteins with molecular weights in the range of 20-25,000. They are stable to acid pH and are active at concentrations of 10^{-12} - 10^{-13} M. Under the proper induction conditions, the genes for F and Le interferons can be activated in a single cell strain, where they are transcribed and translated coordinately (6,7).

Recently, cDNAs for both F and Le interferons were cloned in E. coli (8,9). Complete DNA sequences of the cloned genes indicate that the two interferons have signal peptides and are each composed of 166 amino acids (10,11). They contain a considerable amount (29%) of amino acid sequence homology (12). Despite the close similarity in amino acid sequence, Le and F interferons are completely unrelated antigenically (13). They also differ in their kinetics of development of the antiviral state, antiviral activity on heterologous cells, stability in vivo, and in many of their clinical actions (3,14-16).

Chou and Fasman have shown that it is possible to predict protein secondary structure from an analysis of amino acid sequence (17-19). The Chou-Fasman method has been used on over one hundred proteins (19). Computed percentages of secondary structure agree very well with estimates based on circular dichroism (reviewed in ref. 19). The predictive accuracy for the determination of α-helical and β-pleated sheet residues in individual proteins is generally in the range of 70-80% by comparison with x-ray crystallographic data. However, the precision of the method is somewhat reduced for polypeptides with less than 10% helix or sheet content (19). Thus, the Chou-Fasman method may be presumed to be sufficiently accurate to allow some preliminary generalizations as to the secondary structure of F and Le interferons.

A comparison of the secondary structure of F and Le interferons might indicate what parts of the molecules are likely to be involved in their common biological activities. Since the two molecules are antigenically unrelated,

regions of similarity might be expected to be on the interior of the molecule and regions of difference on the exterior, indicating something about possible three-dimensional conformation as well. This would provide a starting point for researchers interested in producing more active interferons or synthetic analogs with interferon pharmacological activity.

MATERIALS AND METHODS: Amino acid sequences Sequences for F and Le interferons were obtained from analysis of the cloned cDNAs (10,11).

Prediction of protein conformation Amino acid sequences were analyzed for potential sites of secondary structure by determining single residue conformational and positional preferences according to the Chou-Fasman method (17-19). Sequences were processed in a HP3000 computer using a modified version of a program written by Jeffrey Siegel, Pomona College, CA.

RESULTS: Chou-Fasman analysis enabled each amino acid residue of F and Le interferons to be assigned one of the following secondary structural configurations: α helix, β pleated sheet, random coil, or reverse turn. A summary of the results (Table 1) shows that F interferon has a predicted second-

Table 1

Comparison of F and Le interferons according to predicted secondary structure

Calculated values	F	Le
Number of α helical segments	6	8
Amino acids in helixes	59	92
% total of amino acids	36	55
Number of β sheet segments	7	4
Amino acids in sheets	54	26
% total amino acids	33	16
Number of random coil segments	8	10
Amino acids in random coils	25	16
% of total amino acids	15	10
Number of reverse turns	7	8
Amino acids in reverse turns	28	32
% of total amino acids	17	19

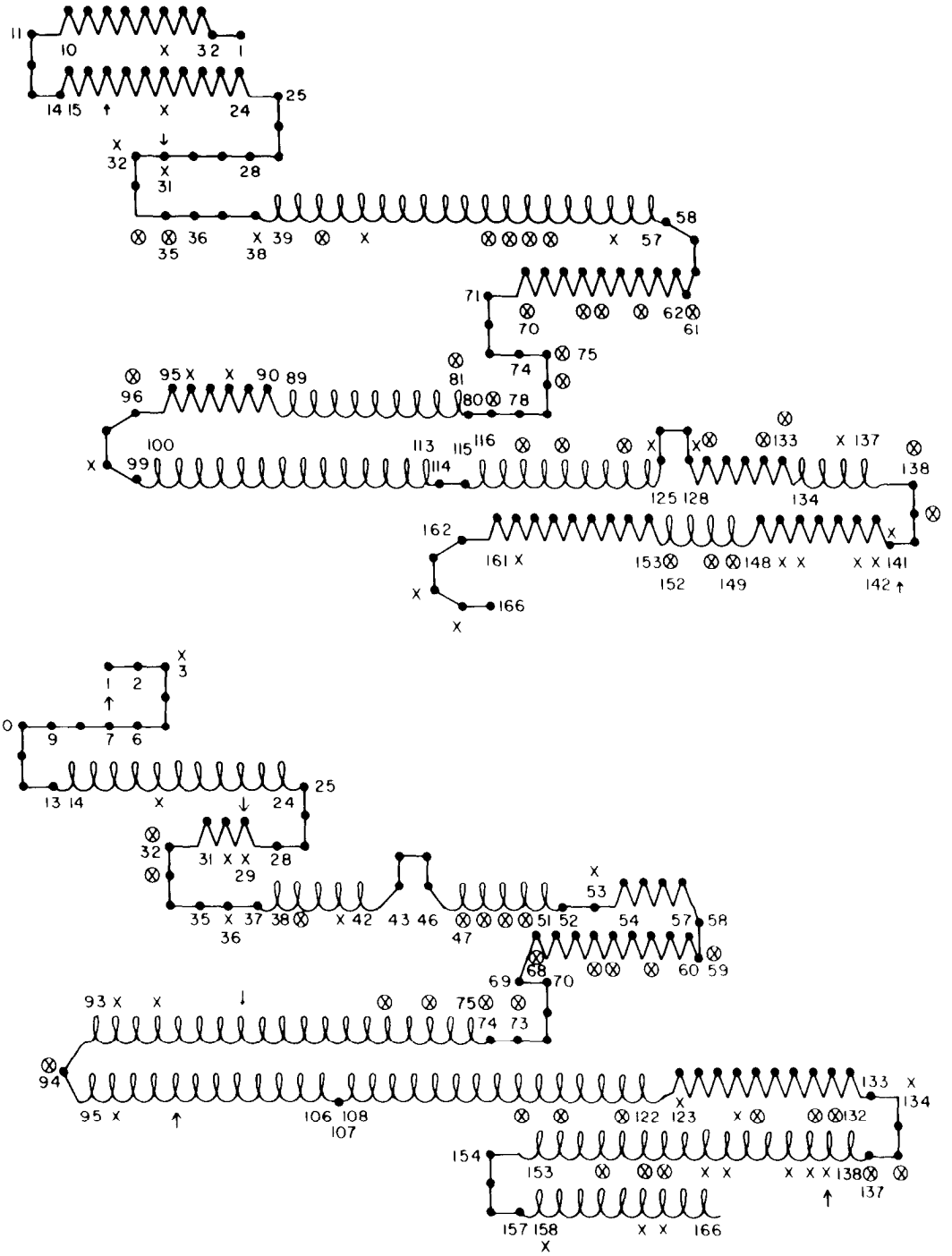

Fig. 1. Diagrammatic sketches of predicted F and Le interferon secondary structures. Each symbol represents one amino acid residue of α helix (ℓ), β -pleated sheet (⋀), or random coil (—). Reverse turns are denoted by U-shaped regions of chain reversal (⊏, ⊓, or ⊐).

upper panel: F interferon lower panel: Le interferon

ary structure containing approximately one-third α helix and one-third β pleated sheet, with the remaining part of the molecule made up equally of reverse turns and random coil. In contrast, over one half of the Le interferon molecule is predicted to be in the α helical conformation, with the rest of the amino acids participating (in decreasing order) in reverse turns, β pleated sheet, and random coil.

Figures 1a and b are schematic diagrams of the secondary structures obtained for F and Le interferons by the Chou-Fasman method of analysis. The molecules have been drawn in such a way as to maximize areas of apparent homology between the two proteins. No attempt has been made to denote 3-dimensional conformation.

It is immediately clear that there are several regions of corresponding secondary structure in F and Le interferons. Most striking is the β pleated sheet at positions 60-68 in Le interferon and 62-70 in F interferon. Table 2 lists amino acid positions predicted to be in the same conformation in the two molecules after adjusting for the observed two amino acid shift in sequence between homologous residues in the two proteins (12). Most strongly conserved between the two proteins are the areas of α helix, in which 45 posi-

Table 2

Regions of homologous secondary structure

Structure	Residue	No. of residues (% of 2° structure)
α Helix	38-42, 47-51, 79-87, 98-106, 108-111, 114-122, 147-150	45 (76% F, 49% Le)
β Sheet	60-68, 127-131	14 (26% F, 54% Le)
Coil	36, 58-59, 94	4 (16% F, 25% Le)
Reverse turn	10-12, 25-26, 32-33, 70-72, 136-137	11 (39% F, 34% Le)
Total		74 (45% of F and Le)

tions have the same conformation in the two molecules. Amino acid residues with homologous secondary structure are more likely to be found in the middle third or C-terminal end of the polypeptide chains than at the N-terminal end. A total of 74 amino acids, or 45% of the molecule, are predicted to assume corresponding conformations in F and Le interferons.

Recently it was shown that 48 positions in the 166 amino acid interferon polypeptides contain the same amino acid residue in F and Le interferons (12). These positions are marked by "x" in Figs. 1a and b. Of these residues, 28 (57%) are predicted to be in an equivalent secondary structure conformation in the two interferons, indicated by "⊠" in Figs. 1a and b.

The positions of cysteine residues are marked with an arrow in the diagrams. It is interesting to note that cys 31 of F interferon and cys 29 of Le interferon are predicted to assume different conformations, despite being found in homologous positions in the amino acid sequence. The same is true of cys 141 of F and cys 139 of Le interferon. It is not known at present whether these cysteines are involved in the active site of the two interferons. However, the presence of an odd number of cysteine residues in both interferons might have suggested the possibility of an enzymatic function involving a sulfhydryl group.

DISCUSSION: Analysis of the amino acid sequence of F and Le interferons by the Chou-Fasman method of determining secondary structure indicates both differences and similarities in the predicted two-dimensional conformations for the two protein molecules. F interferon contains an approximately equal amount of β pleated sheet and α helical residues, while Le interferon has significantly more α helix. F interferon is predicted to have more than twice as many amino acids in sheet conformation than Le interferon. Since β pleated sheets are usually composed of hydrophobic residues and tend to associate with one another via hydrophobic interactions, perhaps this explains the experimental observation that F interferon appears to be more hydrophobic than Le interferon (20).

The two proteins share many regions with homologous secondary structure. These areas tend to be located in the middle and C-terminal end of the protein molecule. It is interesting to speculate on the possible significance of these areas of homology. For instance, the polypeptide stretch from positions 58 to 72 (Le numbering) has 14 out of 15 residues with the same predicted secondary structure in F and Le interferons and is located near a reverse turn, a structure often found at the protein surface (21). A similar situation is found with residues 127-131, which could possibly participate in a 3-dimensional sheet structure with residues 60-68. These regions are good candidates for involvement in the interferon active site. Segment 98-122, with 22 out of 35 amino acids predicted to be in the same conformation in F and Le interferons, might form part of the interferon molecular backbone. The highly conserved residues 47-51 are also likely to play an important role in interferon function.

Since the two interferons are antigenically dissimilar, it is likely that areas of large differences such as the N-terminal ends are found on the exterior of the molecules. As the N termini contain a fairly large amount of hydrophilic residues (10,11), they could be exposed favorably to the solvent. Alternatively, antigenic differences could be primarily determined by differential glycosylation of the two interferons.

ACKNOWLEDGMENTS: I would like to thank Dr. Randall Murphy and Dr. Jan Vilček for support and encouragement. I am grateful to Dr. Murphy and Dr. Ross Smith for use of computer facilities, Dr. Jeffrey Siegel for supplying the original computer program, Dr. Ken Bell for assistance in modifying the program, Dr. Y.K. Yip for valuable discussions, and Ms. Michele Cassano for typing the manuscript. Dr. Tadatsugu Taniguchi and Dr. Charles Weissmann helped immensely by providing interferon sequences and sequence comparisons prior to publication. T.G.H. is a recipient of Medical Scientist Training Grant GM-07308 from the National Institutes of Health.

REFERENCES:

1. Stewart, W.E. II (1979) The Interferon System. Springer-Verlag, Wien-New York.
2. Bloom, B.R. (1980) Nature 284, 593-595.
3. Dunnick, J.K. and Galasso, G.J. (1979) J. Infect. Dis. 139, 109-123.
4. Knight, E. Jr., Hunkapiller, M.W., Korant, B.D., Hardy, R.W.F. and Hood, L.E. (1980) Science 207, 525-526.

5. Zoon, K.C., Smith, M.E., Bridgen, P.J., Anfinsen, C.B., Hunkapiller, M.W., and Hood, L.E. (1980) Science 207, 527-528.
6. Hayes, T.G., Yip, Y.K. and Vilček, J. (1979) Virology 98, 351-363.
7. Pang, R.H.L., Hayes, T.G., and Vilček, J. (1980) (submitted for publication).
8. Taniguchi, T., Sakai, M., Fujii-Kuriyama, Y., Muramatsu, M., Kobayashi, S., and Sudo, T. (1979) Proc. Japan. Acad. 55 Ser. B, 464-469.
9. Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Ecsödi, J., Boll, W., Cantell, K., and Weissmann, C. (1980) Nature 284, 316-320.
10. Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y., and Muramatsu, M. (1980) Gene 10, 11-15.
11. Mantei, N., Schwarzstein, M., Streuli, M., Panem, S., Nagata, S., and Weissmann, C. (1980) Gene 10, 1-10.
12. Taniguchi, T., Mantei, N., Schwarzstein, M., Nagata, S., Muramatsu, M., and Weissmann, C. (1980) Nature, in press.
13. Havell, E.A., Berman, B., Ogburn, C.A., Berg, K., Paucker, K., and Vilček, J. (1975) Proc. Nat. Acad. Sci. USA 72, 2185-2187.
14. Gardner, L.J. and Vilček, J. (1979) J. Gen. Virol. 44, 161-168.
15. Gresser, I., Bandu, M.-T., Brouty-Boyé, D., and Tovey, M. (1974) Nature (London) 251, 543-545.
16. Billiau, A., De Somer, P., Edy, V.G., De Clercq, E., and Heremans, H. (1979) Antimicrob. Agts. Chemother. 16, 56-63.
17. Chou, P.Y. and Fasman, G.D. (1974) Biochemistry 13, 211-221.
18. Chou, P.Y. and Fasman, G.D. (1974) Biochemistry 13, 222-245.
19. Chou, P.Y. and Fasman, G.D. (1978) Adv. Enzymology 47, 45-148.
20. Jankowski, W.J., von Muenchhausen, W., Sulkowski, E., and Carter, W.A. (1976) Biochemistry 15, 5182-5187.
21. Kunz, I.D. (1972) J. Am. Chem. Soc. 94, 4009-4012.