

Primary Structure of Lysozymes from Man and Goose

R. E. CANFIELD, S. KAMMERMAN, J. H. SOBEL & F. J. MORGAN

Department of Medicine, College of Physicians and Surgeons, Columbia University, New York, New York 10032

The primary structure of goose egg white lysozyme seems to be distinct from that of hen egg white and human leukaemia lysozymes.

STUDIES on the structure of hen egg white lysozyme (N-acetylmuramide glycanohydrolase, EC 3.2.1.17) have provided the basis of a detailed proposal for its mechanism of action^{1,2}. Our results indicate that the primary structures of human leukaemia lysozyme and hen egg white lysozyme have considerable homology and that both share some structural features essential to the mechanism of action of hen egg white lysozyme. X-ray crystallographic data (see preceding article by Blake and Swan³) demonstrate that these two enzymes have similar three dimensional structures. Furthermore, a comparison of the primary structure of hen egg white lysozyme and of human leukaemia lysozyme with that of bovine α -lactalbumin suggests that these three proteins shared a common ancestral gene⁴. In contrast, the NH₂-terminal sequence of goose egg white lysozyme (presented here) suggests that the primary structure of this enzyme is distinctly different from both hen and human leukaemia lysozymes.

Human leukaemia lysozyme (HLL) was isolated from the urine of a patient with chronic monocytic leukaemia, who was excreting between 1 and 2 g of lysozyme daily. The methods of purification of HLL from urine were essentially those reported⁵ for goose egg lysozyme, except that the initial batch adsorption step was accomplished by passing the urine over 'Bio-Rex 70'.

The amino-acid sequence of HLL was determined by studying the tryptic, chymotryptic and peptic peptides of reduced, carboxymethyl-HLL (RCM-HLL) using the methods of intact SCM-HLL and positions 30-60 from a peptide obtained from automated Edman degradation⁷ using direct identification of the phenylthiohydantoin derivatives of the amino-acids. Positions 1-46 were determined by a degradation of intact RCM-HLL and positions 30-60 from a peptide obtained by cyanogen bromide cleavage. In suitable cases, the NH₂-terminal amino-acid sequence of tryptic peptides was also determined by automated Edman degradation.

The primary structure of HLL is shown in Fig. 1 together with the primary structures of HEL and of bovine α -lactalbumin (BAL)⁴. The alignment of residues in Fig. 1 illustrates the greatest possible homology among these three proteins. HLL is now shown to have eight half-cystine residues, as does HEL; previous reports had indicated a content of six half-cystines^{8,9}. The sequence proposed here confirms most of the features of an earlier tentative structure¹⁰, and is consistent with an almost complete sequence of human milk lysozyme reported by Jollès¹¹. There are eight positions in which HLL and BAL are identical but different from HEL, six of which occur in the NH₂-terminal half of these molecules. Conversely, there are fourteen positions in which HEL and BAL are identical but different from HLL, eleven of which occur in the

COOH-terminal half of these molecules. There are thirty-six positions in which all three proteins seem to have identical amino-acid residues, including all eight half-cystines. Thus, there are common elements in the primary structures of all three proteins; this confirms the suggestion of Brew *et al.*⁴ that these molecules evolved from a common ancestral gene.

The primary structures of HEL and HLL have seventy-seven identical amino-acid residues. A disulphide link between the half-cystines in positions 6 and 127 has been identified in HLL, and another fragment, isolated from a pepsin digest of the native enzyme, contains the four half-cystines at positions 65, 77, 81 and 95, from which we infer that the remaining half-cystines at positions 30 and 115 are linked by a disulphide bridge. Thus it seems probable, on the basis of this evidence as well as that for the similar three dimensional structures of these enzymes³, that the disulphide bridges in HLL are identical to those reported for HEL¹² and BAL¹³.

Although both lysozymes possess 129 amino-acids, a comparison of their primary structures suggests that the best alignment is achieved by assuming an insertion in HLL near position 48 and a deletion in the region of position 100. Because the extra residue (either glycine or alanine) seems to be added at the turning point of the β -pleated sheet, the three dimensional structure is probably not significantly altered by the insertion. In fact, this is followed by a region that is the most constant among all three proteins, that is, the residues from positions 51-60, where the only acceptable substitutions are leucine for isoleucine and phenylalanine.

We believe that homologous considerations for the primary structure data indicate that a deletion in HLL is located near position 100. The change from valine to arginine in position 100 suggests that the course of the HLL polypeptide chain in this region probably differs from that in HEL. (In HEL the valine side chain is buried in the interior and appears to indent the polypeptide chain inward.) We would expect an arginine in that position in HLL to be directed towards the aqueous exterior and thus straighten the polypeptide chain in that region. This straightening, together with the replacement of glycine at position 102 by proline, may compensate for the deletion. Other implications of the primary structure differences are discussed by Blake and Swan³. A number of these differences occur in the cleft region and may account for the observation that HLL is more active than HEL⁹.

We have determined the NH₂-terminal sequence of goose egg white lysozyme (GEL) isolated⁵ from white Emben geese. Automated Edman degradation was performed on 0.25 μ mol of carboxymethyl-GEL. The amino-acid sequence of the initial thirty residues was found to be:

NH ₂ -	Arg-	Thr-	Asp-	Cys-	Tyr-	Gly-	Asn-	Val-	Asn-
					5				
Arg-	Ile-	Asp-	Thr-	Thr-	Gly-	Ala-	Ser-	Cys-	Lys-
10					15				
Thr-	Ala-	Lys-	Pro-	Glu-	Gly-	Ile-	Ser-	Tyr-	Cys-
20					25				
Gly-									
30									

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
HLL	LYS	VAL	PHE	GLU	ARG	CYS	GLU	LEU	ALA	ARG	THR	LEU	LYS	ARG	LEU	GLY	MET	ASP	GLY	TYR
HEL	LYS	VAL	PHE	GLY	ARG	CYS	GLU	LEU	ALA	ALA	ALA	MET	LYS	ARG	HIS	GLY	LEU	ASP	ASN	TYR
BAL	GLU	GLN	LEU	THR	LYS	CYS	GLU	VAL	PHE	ARG	GLU	LEU	LYS	...	...	ASP	LEU	LYS	GLY	TYR
	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	
HLL	ARG	GLY	ILE	SER	LEU	ALA	ASN	TRP	MET	CYS	LEU	ALA	LYS	TRP	GLU	...	SER	GLY	TYR	ASN
HEL	ARG	GLY	TYR	SER	LEU	GLY	ASN	TRP	VAL	CYS	ALA	ALA	LYS	PHET	GLU	...	SER	ASN	PHE	ASN
BAL	GLY	GLY	VAL	SER	LEU	PRO	GLU	TRP	VAL	CYS	THR	THR	...	PHE	HIS	THR	SER	GLY	TYR	ASP
	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
HLL	THR	ARG	ALA	THR	ASN	TYR	ASN	ALA	GLY	ASP	ARG	SER	THR	ASP	TYR	GLY	ILE	PHE	GLN	ILE
HEL	THR	GLN	ALA	THR	ASN	ARG	ASN	THR	...	ASP	GLY	SER	THR	ASP	TYR	GLY	ILE	LEU	GLN	ILE
BAL	THR	GLU	ALA	ILE	VAL	GLU	ASN	...	...	ASN	GLN	SER	THR	ASP	TYR	GLY	LEU	PHE	GLN	ILE
	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
HLL	ASN	SER	ARG	TYR	TRP	CYS	ASN	ASP	GLY	LYS	THR	PRO	GLY	ALA	VAL	ASN	ALA	CYS	HIS	LEU
HEL	ASN	SER	ARG	TRP	TRP	CYS	ASN	ASP	GLY	ARG	THR	PRO	GLY	SER	ARG	ASN	LEU	CYS	ASN	ILE
BAL	ASN	ASN	LYS	ILE	TRP	CYS	LYS	ASN	ASP	GLN	ASP	PRO	HIS	SER	SER	ASN	ILE	CYS	ASN	ILE
	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
HLL	SER	CYS	SER	ALA	LEU	LEU	GLN	ASP	ASN	ILE	ALA	ASP	ALA	VAL	ALA	CYS	ALA	LYS	ARG	VAL
HEL	PRO	CYS	SER	ALA	LEU	LEU	SER	SER	ASP	ILE	THR	ALA	SER	VAL	ASN	CYS	ALA	LYS	LYS	ILE
BAL	SER	CYS	ASP	LYS	PHE	LEU	ASN	ASN	ASP	LEU	THR	ASN	ASN	ILE	MET	CYS	VAL	LYS	LYS	ILE
	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	
HLL	ARG	...	ASP	PRO	GLN	GLY	ILE	ARG	ALA	TRP	VAL	ALA	TRP	ARG	ASN	ARG	CYS	GLN	ASN	ARG
HEL	VAL	SER	ASP	GLY	ASP	GLY	MET	ASN	ALA	TRP	VAL	ALA	TRP	ARG	ASN	ARG	CYS	LYS	GLY	THR
BAL	LEU	...	ASP	LYS	VAL	GLY	ILE	ASN	TYR	TRP	LEU	ALA	HIS	LYS	ALA	LEU	CYS	SER	GLU	LYS
	119	120	121	122	123	124	125	126	127	128	129									
HLL	ASP	VAL	ARG	GLN	TYR	VAL	GLN	GLY	CYS	GLY	...	VAL								
HEL	ASP	VAL	GLN	ALA	TRP	ILE	ARG	GLY	CYS	ARG	...	LEU								
BAL	LEU	ASP	GLN	...	TRP	LEU	...	...	CYS	LYS	GLU	LEU								

Fig. 1 Amino-acid sequence of human leukaemia lysozyme (HLL) reported here in alignment with the sequence of hen egg white lysozyme (HEL) and bovine α -lactalbumin (BAL) (refs. 4 and 6). The numbers correspond to the sequence of HLL. The HEL and BAL data are arranged to optimize homologous relationships which are indicated by shading. The half-cystines appear in identical locations in all three proteins.

This NH₂-terminal sequence, which represents approximately a quarter of the covalent structure of GEL, is not similar to that of hen egg white lysozyme⁶, the closely related duck lysozymes¹⁴, turkey lysozyme¹⁵, or the HLL reported here. Nor is it similar to the NH₂-terminal sequences of the papaya¹⁶ or the T₄ bacteriophage¹⁷ lysozymes.

These findings confirm earlier suggestions that GEL represents a different class of lysozymes from the known avian lysozymes. GEL is much less stable to heat than HEL¹⁸, is not inhibited by N-acetyl-D-glucosamine or its polymers^{19,20} and is not inactivated by the 2,3'-epoxypropyl- β -glycosides of N-acetyl-D-glucosamine oligomers²¹. Arnheim and Steller²², on the basis of immunological studies of black swan egg white, have noted the presence of two distinct lysozymes which cross-react respectively with antisera to HEL and GEL. They suggest that these two lysozymes may be the product of non-allelic genes specifying two different lysozymes, one "hen-like" and the other "goose-like" in nature. There is not sufficient sequence data for GEL at present to indicate whether this is an example of convergent evolution or, alternatively, of a divergent process.

From the available data we conclude that lysozyme molecules have evolved in nature along several independent genetic lines. Generally, these proteins have evolved as basic molecules that are monomers with molecular weights ranging from 13,000–29,000. Further structural studies of the different classes of lysozymes will probably yield several distinctly different architectural solutions to the problem of facilitating the hydrolysis of glycosidic linkages in bacterial cell walls.

We thank Dr Elliott Osserman for discussions and Dr Colin Blake for X-ray crystallographic data. R. E. C. is a career investigator of the Health Research Council of New York. This study was supported by an NIH research grant.

Received April 5, 1971.

- Phillips, D. C., *Proc. US Nat. Acad. Sci.*, **57**, 484 (1967).
- Johnson, L. N., Phillips, D. C., and Rupley, J. A., *Brookhaven Symp. Biol.*, No. 21, 120 (1968).
- Blake, C. C. F., and Swan, I. D. A., *Nature New Biology*, **232**, 12 (1971).
- Brew, K., Castellino, F. J., Vanaman, T. C., and Hill, R. L., *J. Biol. Chem.*, **245**, 4570 (1970).
- Canfield, R. E., and McMurry, S., *Biochem. Biophys. Res. Commun.*, **26**, 38 (1967).
- Canfield, R. E., *J. Biol. Chem.*, **238**, 2698 (1963).
- Edman, P., and Begg, G., *Europ. J. Biochem.*, **1**, 80 (1967).
- Jollès, J., and Jollès, P., *Biochemistry*, **6**, 411 (1967).
- Osserman, E. F., and Lawlor, D. P., *J. Exp. Med.*, **124**, 921 (1966).
- Canfield, R. E., cited in Johnson, L. N., Phillips, D. C., and Rupley, J. A., *Brookhaven Symp. Biol.*, No. 21, 136 (1968).
- Jollès, P., *Chimica*, **25**, 1 (1971).
- Canfield, R. E., and Liu, A. K., *J. Biol. Chem.*, **240**, 1997 (1965).
- Vanaman, T. C., Brew, K., and Hill, R. L., *J. Biol. Chem.*, **245**, 4583 (1970).
- Hermann, J., and Jollès, J., *Biochim. Biophys. Acta*, **200**, 178 (1970).
- LaRue, J., and Speck, jun., J. C., *J. Biol. Chem.*, **245**, 1985 (1970).
- Howard, J. B., and Glaser, A. N., *J. Biol. Chem.*, **244**, 1399 (1967).
- Inoué, M., Imada, M., and Tsugita, A., *J. Biol. Chem.*, **245**, 3479 (1970).
- Dianoux, A. C., and Jollès, P., *Biochim. Biophys. Acta*, **133**, 472 (1967).
- Jollès, P., Blancard, J. S., Charlemagne, D., Dianoux, A. C., Jollès, J., and Baron, J. L., *Biochim. Biophys. Acta*, **151**, 532 (1968).
- Charlemagne, D., and Jollès, M. P., *CR Acad. Sci.*, **270**, 2721 (1970).
- McKelvy, J., Eshdat, Y., and Sharon, N., *Fed. Proc.*, **29**, 532 (1970).
- Arnheim, N., and Steller, R., *Arch. Biochem. Biophys.*, **141**, 656 (1970).