Expression and disulfide-bond connectivity of the second ligand-binding repeat of the human LDL receptor

Stephan Bieri, Julianne T. Djordjevic, Negar Jamshidi, Ross Smith, Paulus A. Kroon*

Centre for Protein Structure, Function and Engineering, Department of Biochemistry, University of Queensland, Brisbane, Queensland 4072, Australia

Received 12 July 1995; revised version received 2 August 1995

Abstract The human LDL receptor (LDLR) has a binding domain which consists of seven contiguous ligand-binding (LB) repeats, each ~40 amino acids long with three disulfide bonds. The second LB repeat, which is required for full binding of LDL, has been expressed, purified and folded to yield a single, fully oxidized isomer. By selective reduction and alkylation, we have shown that the cysteine residues have a I-III, II-V, IV-VI connectivity, matching that recently determined for the amino-terminal repeat. We suggest that the first two LB repeats of the LDLR, with their unique disulfide-bonding pattern, serve as a structural paradigm for other LB repeats.

Key words: LDL receptor; Peptide; Disulfide; Expression; Mass spectrometry; TCEP

1. Introduction

The LDL receptor (LDLR) is the prototype of a family of cell surface receptors which mediate the endocytosis of multiple ligands [1]. These receptors share a group of common structural elements, which include cysteine-rich ligand-binding (LB) or complement repeats [1]. The LDLR has a binding domain which consists of seven contiguous LB repeats, each ~40 amino acids long [2], with a repeating pattern of six cysteine residues and a highly conserved series of negatively charged amino acids near the COOH-terminus. The LDLR utilizes these repeats to bind plasma lipoproteins that contain apolipoprotein (apo) B-100 [3] and apoE [4]. Binding is believed to involve ionic interactions between acidic residues in the repeats and clusters of basic amino acids in apoB-100 and apoE [2]. LB repeats are not restricted to the LDL receptor family of cell surface receptors, but are found in diverse proteins, including complement factors [5], the proteoglycan perlecan [6], a receptor for subgroup A Rous Sarcoma Virus [7] and the linker chain of the hemoglobin of *Lumbricus terrestris* [8].

As a first step towards understanding the structure and function of LB repeats, we have recently expressed the amino-terminal repeat (rLB1) of the LDLR and determined its disulfide connectivity (S. Bieri, unpubl. obs.) and 3-dimensional structure [9]. This repeat was chosen because of the availability of a conformation-specific monoclonal antibody (IgG-C7) which provided a means of analysing the conformation of the expressed and folded repeat [10]. rLB1 folded into a single isomer with three unique disulfide bonds, connecting Cys(I)–Cys(III), Cys(II)–Cys(V) and Cys(IV)–Cys(VI)[1]. However, LB1 is unique amongst the LB repeats, in that it is not required for ligand-binding. To determine whether the unique disulfide-bonding pattern described for rLB1 is conserved in other LB repeats, we have extended our studies to the expression, folding and disulfide-bonding pattern for the second LB repeat, LB2. We chose this repeat because it plays a role in the binding of apoB-100 [11,12]. Patients with a splicing defect, which leads to the deletion of the exon which encodes LB2, develop a severe clinical phenotype of familial hypercholesterolemia, which has been attributed in part to a reduced affinity of the mutant receptor for LDL [13].

2. Materials and methods

Vinylpyridine, bovine thrombin, iodoacetamide, GSH-agarose and sequencing-grade trypsin were obtained from Sigma. Tris-(2-carboxyethyl)-phosphine (TCEP) was kindly provided by Alun Jones (University of Queensland). Other chemicals used were reagent grade.

2.1 Expression, folding and purification of rLB2

The cDNA-encoding LB2 (residues 43–83 of the mature LDLR) was PCR-amplified using primers which contained *Eco*RI and *Bam*HI sites (primer 1: GAGGGATCCTTGTCTGTCACCTGC; primer 2: CCGAATTCTCAACAGCCTTGCTCGTCTGA), digested with *Eco*RI and *Bam*HI and cloned into the expression vector pGEX-2T, to give pGST-LB2. Recombinant LB2 (rLB2) was expressed in *E coli DH5α* using a 15-l Chemap fermenter (Chemap, Männedorf, Switzerland) and was purified as previously described [14]. rLB2 (250–400 μg/ml) was folded overnight in presence of 3 mM GSH/0.3 mM GSSG in 50 mM Tris-HCl, 150 mM NaCl, 2.5 mM $CaCl_2$, pH 8.5, at 4°C, and was purified using an Econosil C18 reverse-phase HPLC column (10-μm particle size, 60-Å pore, 10 × 250 mm; Alltech, Deerfield, IL). A trifluoroacetic acid/acetonitrile solvent system (solvent A: 0.1% trifluoroacetic acid in water; solvent B: 0.1% trifluoroacetic acid in acetonitrile) was used for the purification, typically with a flow rate of 2.5 ml/min. FPLC analyses were performed on a Smart System (AMRAD-Pharmacia, Australia), using a μRPC C2/C18 reverse-phase column (3-μm particle size, 120-Å pore, 2.1 × 100 mm) with a trifluoroacetic acid/acetonitrile buffer system (see above).

2.2 Reduction and alkylation with iodoacetamide

Partial reduction and alkylation of rLB2 was performed as previously described [15]. Briefly, HPLC-purified rLB2 (2 μg in 15–25 μl of elution buffer) was mixed with an equal volume of 20 mM Tris-(2-carboxyethyl)-phosphine (TCEP) in 20 mM ammonium carbonate, pH 3, and incubated for 7–10 min at 10°C The partially reduced peptide was immediately alkylated with supersaturated iodoacetamide (~2.2 M), quenched ~25 s later with 200 μl of 0.5 M citric acid and injected

*Corresponding author. Fax: (61) (7) 3365-4699.
E-mail: pkroon@biosci.uq.oz.au

Abbreviations: LDL, low-density lipoprotein; LDLR, LDL receptor; LB, ligand-binding; LB1 (LB2), first (second) LB repeat; rLB1 (rLB2), recombinant LB1 (LB2); RP, reverse-phase, TCEP, Tris-(2-carboxyethyl)-phosphine; AM, acetamide; PE, pyridylethyl; AM_n-rLB2, rLB2 with *n* acetamide-labelled cysteine residues; $AM_n/PE_{(6-n)}$-rLB2, rLB2 with *n* acetamide- and (6−*n*) pyridylethyl-labelled cysteine residues.

[1] Cysteine residues identified with Roman numerals (Cys(I) through Cys(VI)) refer to the order of cysteine residues and not their absolute position.

SSDI 0014-5793(95)00939-6

onto a RP-FPLC column. Different alkylated forms of rLB2 (AM_n-rLB2), where n corresponds to the number of acetamide (AM) groups incorporated, were separated using a 24–36% acetonitrile gradient over 24 min.

2.3. Reduction and alkylation with vinylpyridine

AM_n-rLB2 (0.5 μg in 50–100 μl of FPLC elution buffer) was incubated with 100 μl of 0.5 M Tris acetate, pH 8, 2 mM Na_2EDTA, 20 mM DTT, at 65°C for 15 min. Following addition of 5 μl of 4-vinylpyridine, the reaction mixture was incubated in the dark, at room temperature, for 20 min. The reaction was stopped by adding 100 μl of 0.5 M citric acid. This resulted in the formation of rLB2 which was labelled with both AM and pyridylethyl (PE) groups ($AM_n/PE_{(6-n)}$-rLB2) The peptide was purified by RP-FPLC using a 24–36% acetonitrile gradient over 24 min.

2 4. Tryptic digestion of AM/PE-labelled rLB2

Peptide (2–5 μg) was taken to dryness in a Savant SpeedVac SC100, dissolved in 100 μl of 50 mM NH_4HCO_3, pH 8.5, containing 2 mM $CaCl_2$ and cleaved with trypsin using a 1:10 (w/w) ratio of enzyme to protein, for 16 h at 37°C. Tryptic fragments were separated by RP-FPLC using a 1–30% acetonitrile gradient over 60 min and individual peaks collected.

2.5 Mass spectrometry and amino acid sequence analysis

Peptides were analysed using a triple quadrupole mass spectrometer (SCIEX API III, Thornhill, Canada), equipped with an ionspray interface and operated in the positive detection mode. Data were analysed using SCIEX software (MacSpec version 3.2). Peptides (0.1–1.0 nmol) were sequenced using automated Edman degradation on an Applied Biosystems 470A protein gas-phase sequencer with on-line PTH analysis

3. Results and discussion

LB2 was expressed as a thrombin-cleavable GST fusion protein in *E. coli*, using the expression vector pGST-LB2 as described in Section 2. The fusion protein was bound to GSH-agarose and cleaved with thrombin to release the 43 amino acid peptide rLB2:

```
43          53          63          73          83
.           .           .           .           .
GSLSVTCKSGDFSCGGRVNRCIPQFWCDGQVDCDNGSDEQGC
```

The first two amino acids of this peptide (GS) are derived from the thrombin cleavage site [14] and are followed by residues 43–83 of the mature LDL receptor. The peptide was folded in the presence of a GSH/GSSG redox buffer at 4°C for at least 12 h, to yield a form which eluted as a single peak from a RP-FPLC column (Fig. 1), consistent with the presence of a single isomer. This isomer was purified by preparative RP-HPLC. Folded and RP-HPLC-purified rLB2 was used for all subsequent studies described below.

Table 1
Analysis of peptides formed from partial reduction and alkylation of rLB2[a]

Peptide[b]	M_r	n[c]
1	4932	6
2	4816	4
3	4582	0
4	4816	4
5	4816	4
6	4701	2
7	4700	2

[a]Peptides were produced by partial reduction and alkylation of rLB2, purified by RP-FPLC as described in Fig 2 and were analysed by mass spectrometry.
[b]Peptide numbers correspond to peaks in Fig. 2. Peptide peaks were combined from several RP-FPLC runs.
[c]The number of alkylated cysteine residues (N) in each of the isolated components was calculated using M_r = 4582 for the completely oxidized peptide (peptide 3) and a mass of 58 Da for the AM group.

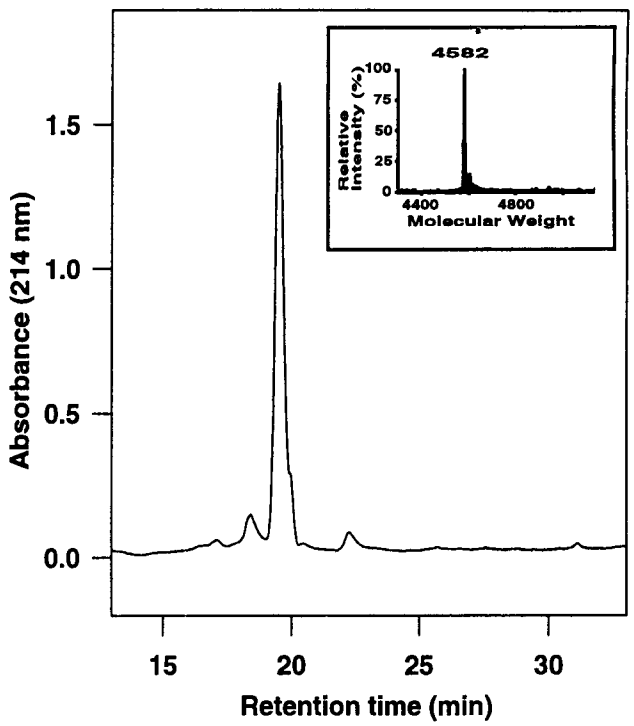

Fig 1. RP-FPLC analysis of rLB2 following thiol-disulfide exchange. A 10–40% acetonitrile gradient over 30 min was used. Inset: mass spectral analysis of rLB2 following incubation in 6 M GnHCl (2 h at 50°C) and injection into 2.2 M iodoacetamide. The treated peptide was immediately purified by RP-FPLC and analysed by mass spectrometry.

Mass spectral analysis of refolded rLB2 gave a molecular mass of 4582 ± 1 mass unit, consistent with the presence of 3 disulfide bridges. Calculated values for reduced and fully oxidized forms of rLB2 are 4588 and 4582. To confirm the complete absence of free cysteine-thiol groups, rLB2 was incubated in the presence of 6 M GnHCl for 2 h at 50°C and was rapidly injected into a supersaturated solution of iodoacetamide to alkylate any free thiol groups. Subsequent mass spectral analysis of the peptide showed that its molecular mass was unchanged, with no evidence for species with alkylated cysteine residues (Fig. 1, inset). To confirm that cysteine residues could be completely alkylated under these conditions, rLB2 was reduced with 45 mM DTT prior to the denaturation and alkylation reactions. This resulted in a molecular mass of 4931 ± 1, consistent with a calculated mass of 4930 for rLB2 with 6 alkylated cysteine residues (data not shown). We conclude that the refolded rLB2 domain is completely oxidized and contains no free cysteine thiol groups.

The traditional strategy for analysing disulfide connectivities is to cleave the peptide chain with proteolytic enzymes and to isolate bridged fragments for amino acid analysis. This approach is often unsuccessful for peptides which contain multiple disulfide bonds [15]. Indeed, both rLB1 and rLB2 are resistant to cleavage by trypsin. We used an alternative method in

which disulfide bonds are sequentially reduced at acidic pH with organic phosphines followed by rapid alkylation of resulting free thiol groups with iodoacetamide [15]. In this work we reduced rLB2 with TCEP at pH 3. The abundance of individual reduced peptides formed depended on the temperature and time of the TCEP reaction. For rLB2, reduction at 10°C for 7–10 min resulted in a significant depletion of fully oxidized rLB2 and minimized the formation of fully reduced rLB2. RP-FPLC (Fig. 2) and mass spectral analysis (Table 1) of partially reduced and AM-labelled rLB2 lead to the identification of 7 species, out of a total of 8 theoretically possible species. Molecular masses of peptides 1–7 (derived from peaks 1–7 in Fig. 2) were consistent with species which contained 0, 2, 4 or 6 AM groups (Table 1), derived from peptides in which 0, 1, 2 or 3 disulfide bonds were reduced, respectively. Peak 7 always had a shoulder ('S' in Fig. 2), suggesting that this might contain the 'missing' 8th peptide. Since this peptide, like peptides 6 and 7, was predicted to contain two acetamide groups, mass spectral analysis could not be used to confirm its presence directly.

To identify a single disulfide bond uniquely, partially reduced peptides are required in which either a single disulfide bond has been reduced, or in which a single disulfide bond remains. Because the most prominent peaks which contain such peptides are peaks 4 and 7, the corresponding peptides were used for subsequent analysis. Peptide 7 (AM_2-rLB2) was isolated, taking care to avoid material in the shoulder fraction, and was reduced in the presence of DTT. Free cysteine thiol groups were alkylated with 4-vinylpyridine to yield the doubly labelled peptide AM_2/PE_4-rLB2. Amino acid sequence analysis of residues 1–22 of this peptide showed that both Cys(I) and Cys(III) were AM-labelled, while Cys(II) was PE-labelled. Since peptide 7 contained only two AM-labelled cysteine residues, we conclude that these cysteine residues were disulfide linked in the fully oxidized peptide. These data therefore provide direct evidence for a Cys(I)–Cys(III) disulfide bridge.

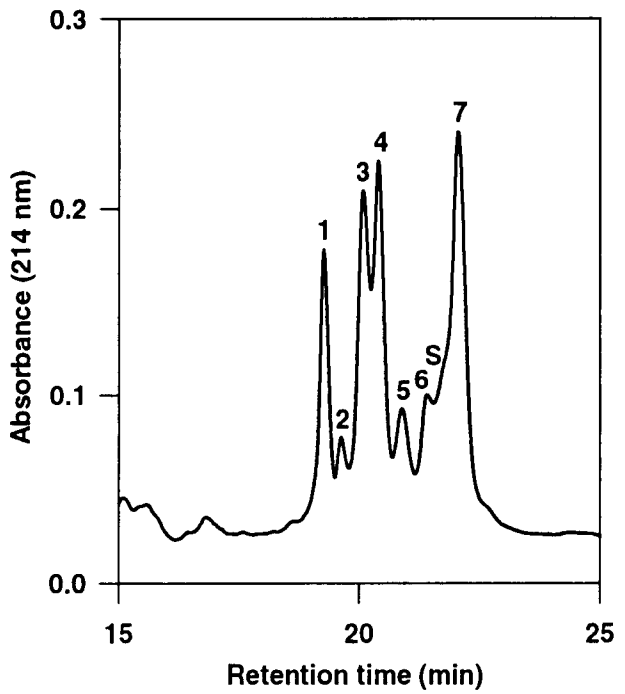

Fig. 2. RP-FPLC separation of partially reduced and alkylated rLB2. Disulfide bonds were partially reduced with TCEP (10 min, 10°C) and the resulting free thiols alkylated by injection into 2 2 M iodoacetamide Resulting peptides were separated by RP-FPLC, using a linear (24–36%) gradient of acetonitrile

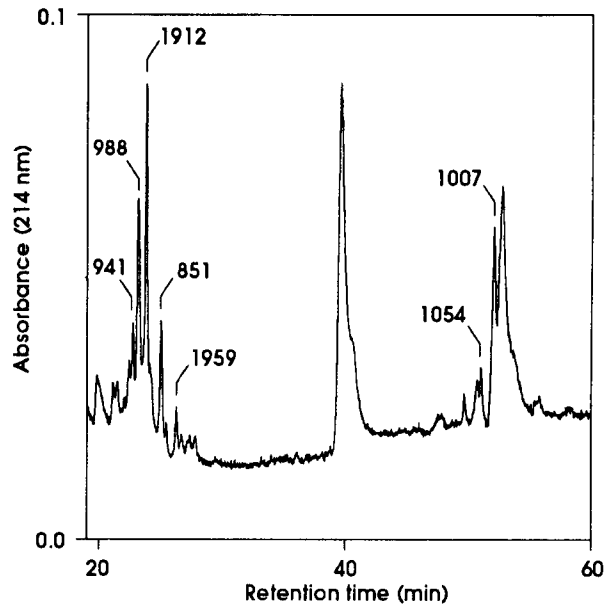

Fig. 3. RP-FPLC separation of a tryptic digest of PE-labelled peptide 7. Peptide 7 was reduced with DTT and alkylated with vinylpyridine to yield the doubly labelled peptide AM_2/PE_4-rLB2. This peptide was digested with trypsin (1:10 w/w) for 16 h at 37°C Cleavage products were separated by RP-FPLC using a 1–30% acetonitrile gradient over 60 min.

Peptide 4 (AM_4-rLB2) was isolated by collecting only half of peak 4 (Fig. 2) in order to minimize contamination by peptide 3 (fully oxidized rLB2). The two cysteine residues in the remaining disulfide bond were labelled with vinylpyridine, following reduction with DTT. The resulting AM_4/PE_2-rLB2 was cleaved with trypsin and cleavage products separated by RP-FPLC (Fig. 3). The locations of tryptic sites (K8, R17, R20 and R27) relative to Cys residues, together with the predicted masses of tryptic fragments, are illustrated in Fig. 4. Mass spectral analysis showed that major fragments were derived from PE_2-labelled peptide-4 (black boxes in Fig. 4), in which both

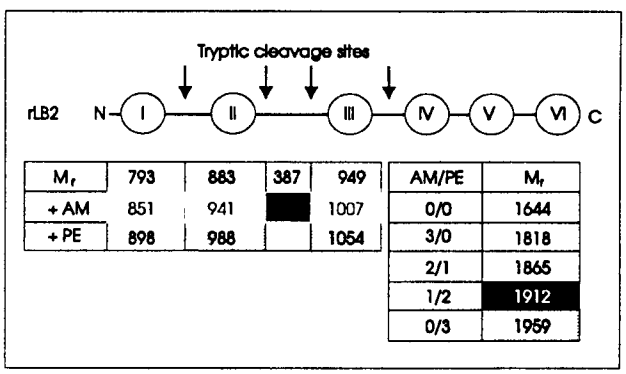

Fig. 4. Identification of peptides formed by tryptic digestion of AM_2/PE_4-rLB2. A schematic representation of rLB2, its tryptic cleavage sites (arrows) and the location of cysteine residues relative to these sites is shown at the top of the figure. The molecular masses of each of the unmodified tryptic peptides, as well as the predicted masses for all combinations of AM/PE-modified fragments, are shown below the representation of the tryptic peptides. Molecular masses for tryptic fragments of peptide 4 are shown in black boxes, while the gray boxes represent cleavage products of peptide 3

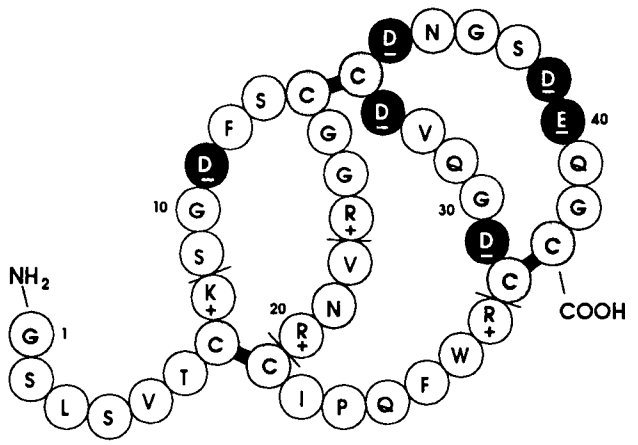

Fig. 5. Structural model for rLB2, which illustrates the tricyclic structure of the disulfide bridged peptide The location of tryptic cleavage sites is indicated by thin lines.

PE-labelled cysteine residues were located in a single fragment comprising Cys(IV)-Cys(VI). However, fragments derived from peptide 3 (PE_6-rLB2 from peak 3 in Fig. 2) were present as minor contaminants (gray boxes in Fig. 4). Indeed, all but one of the predicted fragments derived from peptide 3 were identified. Given the fact that peaks 3 and 4 overlap significantly, the presence of these fragments was not unexpected.

The M_r = 1912 fragment (residues 28–43) which contained Cys(IV), Cys(V) and Cys(VI), was isolated by RP-FPLC chromatography and sequenced. Analysis of the first 8 residues of this fragment (C28-C43) showed that there was minor contamination from the M_r = 851 (residues G1-K8) and M_r = 988 fragments (S9-R17) derived from peaks which flank the M_r = 1912 peak. However, the assignment of residues C28-D35, and in particular the demonstration that Cys(IV) was labelled with a PE group, and that Cys(V) was labelled with an AM group, was unequivocal. Since the Cys(IV)–Cys(VI) fragment contained two PE-labelled cysteine residues, we conclude that Cys(VI) was also PE-labelled and that the two PE-labelled cysteine residues (Cys(IV) and Cys(VI)) were disulfide-bonded prior to reduction and pyridylethylation of peptide 4. In view of the Cys(I)–Cys(III) and Cys(IV)–Cys(VI) connectivities, we infer that the remaining disulfide bond bridges Cys(II) and Cys(V). The tricyclic structure of rLB2 formed by these overlapping disulfide bonds is illustrated in Fig. 5.

The Cys(I)–Cys(III), Cys(II)–Cys(V), Cys(IV)–Cys(VI) connectivity is identical with that found for rLB1 (S. Bieri et al., unpubl. obs.; [9]). We conclude that the same arrangement of disulfide bonds results in thermodynamically favored structures for both rLB1 and rLB2. For rLB1 the resulting structure is immunologically indistinguishable from the same repeat in the LDLR (S. Bieri et al., unpubl. obs.; [9]). Although there are no corresponding antibodies for LB2, the fact that it folds into a single isomer with a disulfide-bonding pattern which is identical with that of rLB1, strongly suggests that this pattern is a general feature of LB repeats.

The 3-dimensional structure of rLB1 contains a β-hairpin following a series of turns [9]. This results in a structure in which most of the acidic residues, including the highly conserved SDE triad, are clustered on one face of the molecule, forming an appropriate charged surface for interaction with basic residues of apoB-100 and apoE. A similar clustering of charged residues is present in rLB2 (N. Daly et al., unpubl. obs.). Despite the apparent structural similarity, LB1 plays no role in the binding of ligands, while LB2 does. Experimentally, one of the major differences between rLB1 and rLB2 is the susceptibility to TCEP- mediated reduction. While a temperature of 65°C was used for he partial reduction of rLB1 (S. Bieri et al., unpubl. obs.), a much lower temperature of 10°C was required for rLB2. We surmise that this is due to a more flexible structure for rLB2, which allows the TCEP molecules greater access to the disulfide bonds. This is supported by [^{1}H]NMR spectroscopic studies, which show that rLB2 is much more flexible than rLB1 (N. Daly et al., unpubl. obs.). It is possible that the greater flexibility of LB repeats is required to allow acidic residues to acquire an appropriate conformation for their interaction with clusters of basic residues on apoB-100 and apoE. Alternatively, specific interactions between different LB repeats may be required to place them in an appropriate environment which allow interactions with either apoB-100 or apoE. Such interactions may depend on a structural element not present in LB1.

Acknowledgements: We would like to thank Ms. N. Daly (University of Queensland) for unpublished information on the 3-dimensional structure of rLB2. This work was supported by a research grant from the National Health and Medical Research Council of Australia to P.A. Kroon and R. Smith.

References

[1] Hobbs, H.H., Russell, D.W., Brown, M.S. and Goldstein, J.L. (1990) Annu. Rev. Genet. 24, 133–170.
[2] Goldstein, J.L., Brown, M.S., Anderson, R.G.W., Russell, D.W. and Schneider, W.J. (1985) Annu. Rev. Cell Biol. 1, 1–39.
[3] Brown, M.S. and Goldstein, J.L. (1986) Science 232, 34–47.
[4] Mahley, R.W. (1988) Science 240, 622–630.
[5] Haefliger, J.A., Tschopp, J., Nardelli, D., Wahli, W., Kocher, H.P., Tosi, M. and Stanley, K.K. (1987) Biochemistry 26, 3551–3556.
[6] Noonan, D.M., Fulle, A., Valente, P., Cai, S., Horigan, E., Sasaki, M., Yamada, Y. and Hassell, J.R. (1991) J. Biol. Chem. 266, 22939–22947.
[7] Bates, P., Young, J.A.T. and Varmus, H.E. (1993) Cell 74, 1043–1051.
[8] Suzuki, T. and Riggs, A.F. (1993) J. Biol. Chem. 268, 13548–13555.
[9] Daly, N., Scanlon, M.J., Djordjevic, J.T., Kroon, P.A. and Smith, R. (1995) Proc. Natl. Acad. Sci. USA (in press).
[10] van Driel, I.R., Goldstein, J.L., Südhof, T.C. and Brown, M.S. (1987) J. Biol. Chem. 262, 17443–17449.
[11] Esser, V., Limbird, L.E., Brown, M.S., Goldstein, J.L. and Russell, D.W. (1988) J. Biol. Chem. 263, 13282–13290.
[12] Russell, D.W., Brown, M.S. and Goldstein, J.L. (1989) J. Biol. Chem. 264, 21682–21688.
[13] Sun, X.-M., Patel, D.D., Bhatnagar, D., Knight, B.L. and Soutar, A.K. (1995) Arterioscler. Thromb. Vasc. Biol. 15, 219–227.
[14] Smith, D.B. and Johnson, K.S. (1988) Gene 67, 31–40.
[15] Gray, W.R. (1993) Prot. Sci. 2, 1732–1748.