Mistletoe lectin I forms a double trefoil structure

Edel C. Sweeney[a], Alexander G. Tonevitsky[e], Rex A. Palmer[a,*], Hidie Niwa[a], Uwe Pfueller[b], Juergen Eck[c], Hans Lentzen[d], Igor I. Agapov[e], Mikhail P. Kirpichnikov[e]

[a]*Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, UK*
[b]*Universität Witten/Herdecke, Institut für Phytochemie, Stockumer Straße 10, 58453 Witten, Germany*
[c]*Biotechnology Research and Information Network (BRAIN) GmbH, Darmstadter Straße 34, 64673 Zwingenberg, Germany*
[d]*Madaus AG, Osterheimer Straße 198, D-51109 Cologne, Germany*
[e]*Centre of Bioengineering of the Russian Academy of Sciences, pr. 60 let Oktyabrya 7/1, Moscow, Russia*

Received 10 June 1998

Abstract The quaternary structure of mistletoe lectin I (MLI), a type II ribosome inactivating protein, has been determined by X-ray crystallography. A definitive molecular replacement solution was determined for MLI using the co-ordinates of the homologue ricin as a search model. MLI exists as an $[AB]_2$ dimer with internal crystallographic two-fold symmetry. Domain I of the B chains is non-covalently associated through interactions involving three looped chains (α, β, γ) in each molecule of the dimer, forming a double trefoil structure. The ricin molecule which shares 52% sequence homology with MLI has a disulphide bridge between Cys^{20} and Cys^{39} in the α loop. An evolutionary mutation has replaced Cys^{39} with serine in MLI. This mutation appears to allow the α loop the flexibility required to take up its place at the dimer interface, and also suggests a rationale for why ricin does not form dimers. Measurement of retention times using FPLC gel filtration confirms that dimerisation also occurs in solution between MLI B chains with an association constant $K_a = 10^6$ M.

Key words: Mistletoe lectin; Viscumin; Amino acid sequence; X-ray crystallography; Trefoil; Ricin; Ribosome inactivating protein; Dimerisation

1. Introduction

Mistletoe lectins from *Viscum album* are type II ribosome inactivating proteins (RIPs), comprising a toxic A chain and carbohydrate binding B chain linked by a disulphide bond, like other members of this family, including ricin, abrin and modecin [1]. Although these toxic proteins are of different phylogenetic origin they share similar structural properties and enzyme activity, being potent inhibitors of eukaryotic protein synthesis at the ribosomal level. The A chain inhibits protein synthesis by enzymatic depurination of an adenosine residue (A4324) in a highly conserved RNA loop of the 28S ribosomal subunit [2]. Endocytosis of the protein is mediated by the B chain lectin which has multiple potential sugar binding sites, interacting with glycoproteins on the surface of target cells. Following endocytosis the exact process of intracellular trafficking is unclear but current work implicates the endoplasmic reticulum in translocation of the A chain (for a review see Lord and Roberts [3]). Three toxic lectins, MLI, MLII and MLIII isolated from the mistletoe plant, share a common primary structure homology as has been confirmed by N-terminal sequencing of the three A chains [4]. Molecular weight analysis by gel filtration indicates that MLI (115 kDa) is the only member of the RIP type II mistletoe family in its propensity to associate to form non-covalent dimers, while MLII (64 kDa) and MLIII (61 kDa) remains as monomers [5]. However, these proteins show different levels of affinity for galactose and *N*-acetylgalactosamine and their cytotoxic effects on target cells also differ. The biochemical properties of RIPs have been well documented although the biological function remains unclear. The ribosome inactivating properties of the A chain suggest a role in plant defence against bacterial or fungal attack. Some type II RIPs including MLI, MLIII, abrin agglutinin and ricin also exhibit cryoprotective properties, protecting plant cell membranes from freeze-thaw damage by binding to galactolipid head groups. Once binding is achieved, hydrophobic interactions between membrane and protein determine the extent of cryoprotection [6].

2. Materials and methods

MLI has been cloned and the amino acid sequence of prepro MLI was determined from the gene (Patent EP0 751 221 A1, 26.06.1995. Rekombinantes Mistellektin (rML)). Using the CLUSTAL V alignment software [7], sequence comparison with ricin was performed with PAM250 homology matrix for achieving maximum identity. MLI A is comprised of 252 residues and is 41% homologous to ricin A chain (RTA). MLI B contains 263 residues and exhibits 63% homology with ricin B chain (RTB). A 16 residue linker peptide connects both chains in the prepro form of the protein. To produce the full length sequence of the mistletoe gene several independent PCR approaches were performed starting from genomic plant DNA.

MLI, RCA, ricin and their corresponding subunits were purified as described previously [8].

Retention times of MLI and its individual A chain and B chain in solution were assessed. Ricin was used as a control. Retention time analysis was performed on a Superdex G.75 column (10 mm×300 mm). A flow rate of 1 ml/min was used, and the elution buffer contained sodium phosphate buffer, pH 6.8. The amount of injected protein was 100 μl at a concentration of 50 μg/ml. The retention time of each protein was recorded, in order to determine whether or not dimerisation had occurred. These experiments were performed in the presence of lactose to prevent affinity retention. As shown in Table 1 ricin has a retention time of 72.0 s corresponding to its [AB] monomeric state in solution. MLI in contrast has a significantly lower retention time of 61.5 s, confirming the presence of an $[AB]_2$ dimer

*Corresponding author. Fax: (44) (171) 631 6804.

Abbreviations: ML, mistletoe lectin; MLI A, mistletoe lectin I A chain; MLI B, mistletoe lectin I B chain; RTA, ricin A chain; RTB, ricin B chain; RIP, ribosome inactivating protein; R_{merge}, $\Sigma_h\Sigma_i[\|F_o|^2{}_i-(|F_o|^2)_{av})_h]/\Sigma_h\Sigma_i(|F_o|^2)_{av})_h$, provides an overall agreement index between X-ray reflections and symmetry equivalents measured more than once; *R*-factor, $\Sigma\||F_o|-|F_c\||/\Sigma|F_o|$ summed over unique hkl data; Correlation coefficient, $[\Sigma_h[\Delta|F_o|\Delta|F_c|]\times\{[\Sigma_h(\Delta|F_o|)^2]\times[\Sigma_h(\Delta|F_c|)^2]\}^{-1/2}$ where $\Delta|F_h|$ stands for $|F_h|-|F_k|_{av}$ corresponding to the highest peak of the molecular replacement translation function

PII: S0014-5793(98)00766-2

in solution. Retention times of individual chains confirmed the crystallographic finding that dimerisation occurs between B chains. MLI B has a significantly lower retention time of 69.0 s compared to RTB of 81.0 s. RTA and MLI A retention times of 79.5 s each are consistent with the existence of single chain monomers.

MLI crystals were grown from sodium phosphate buffer pH 6.7, containing 50% saturated ammonium sulphate, as previously described [9]. X-ray intensity data were collected on a 30 cm Mar Research Image Plate detector using a synchrotron radiation source, $\lambda = 0.92$ Å. The crystal to detector distance was set at 450 mm in order to resolve reflections along the c* axis. A total of 130.5° of X-ray data was collected from two crystals to approximately 3.7 Å resolution. These data were indexed using DENZO software [10]. The hexagonal unit parameters refined to $a = b = 110.79$ Å $c = 308.53$ Å, in space group $P6_522$, with a single MLI monomer per asymmetric unit, including 60% solvent. The X-ray intensity data were scaled and merged using the CCP4 suite of programmes [11]. The total number of reflexions accepted was 64 381 which reduced to 11 241 unique reflexions with an R_{merge} of 0.09 for $d_{min} = 3.7$ Å. The ricin structure previously described [12], Protein Data Bank ID code 2AAI, yielded a definitive molecular replacement solution employing the CCP4 version of AMoRe [13]. The search model Patterson maps were calculated in the resolution range 3.7–20 Å by placing the ricin structure into an orthogonal cell of P1 symmetry with unit cell parameters 120 Å × 100 Å × 80 Å. The rotation function was stepped over 2.5° and the radii of integration varied between 15 Å and 30 Å. The overall temperature factor was set to −20 Å^2 to sharpen the data. In space group $P6_522$ a translation function solution was found with a peak height of 5.8 σ above the highest noise peak with an *R*-factor of 46% and a correlation coefficient of 52% following the rigid body refinement protocol of AMoRe. No molecular replacement solution was found for the enantiomorphic space group $P6_122$. All symmetry related molecules in the unit cell were generated and examined graphically using MOLPACK [14]. No stereochemical clashes were observed and the packing of the molecule in the crystal clearly demonstrates the interface region between MLI B chain and its neighbouring symmetry equivalent. Electron density maps ($|F_o|-|F_c|$) and ($2|F_o|-|F_c|$) were computed within the resolution range 3.7–10.0 Å by using phases calculated from the ricin model and displayed using the graphics program O [15]. The initial ($2|F_o|-|F_c|$) electron density map contoured to 1.25 times the RMS density level showed good connectivity for the polypeptide chains. The sequence of MLI was fitted to the structure using electron density as a guide and local refinement.

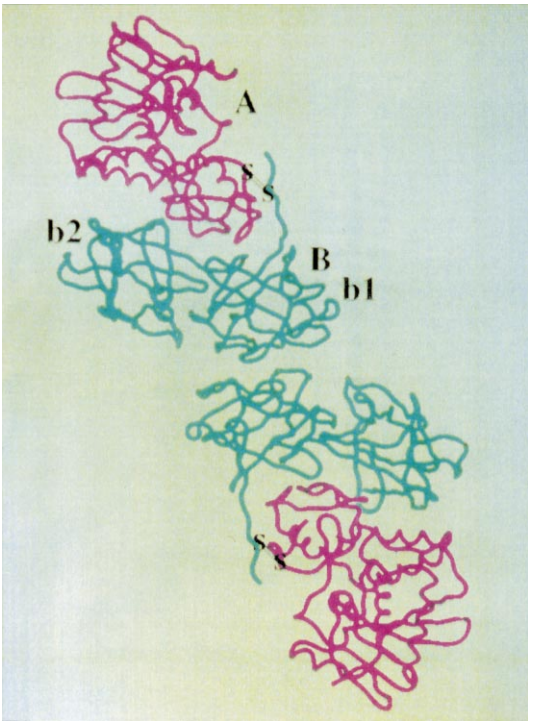

Fig. 1. MLI dimer. A chains (A) are pink, B chains (B) are blue. Domain 1 and domain 2 of the B chain are indicated as b1 and b2. The disulphide bridge between the A and B chains is indicated.

3. Results and discussion

The structure of MLI in the crystal lattice was examined to determine structural features which accommodate its characteristic homodimerisation. The two molecules in the dimer related to each other by the crystallographic two-fold axis

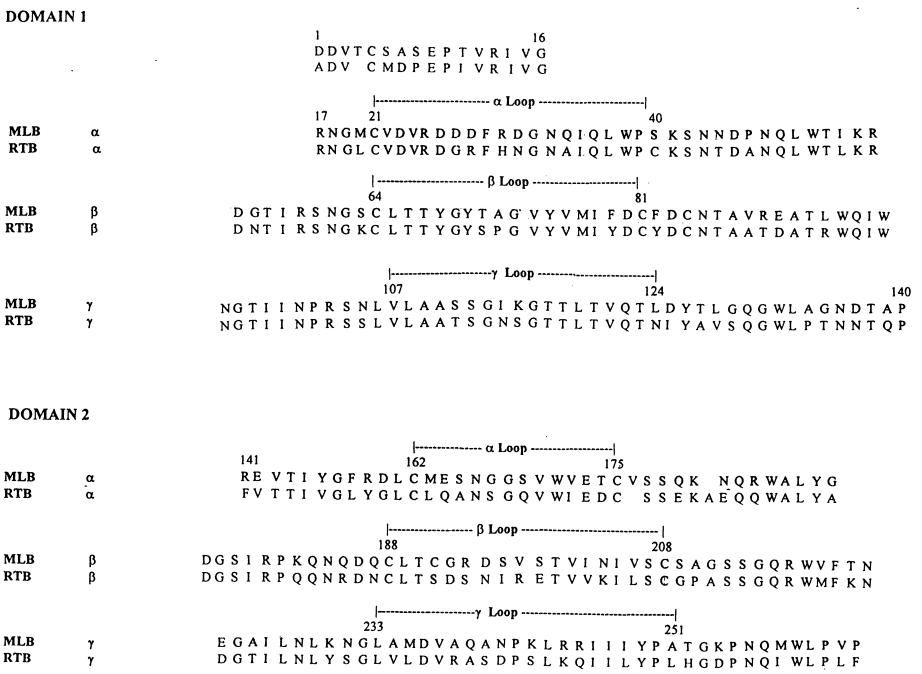

Fig. 2. Sequence alignment of ricin and MLI B chain domain 1. The three hairpins or loops are indicated as α, β and γ.

Table 1
Retention times for MLI proteins, using ricin as a control

Sample injected 100 μl	Retention times (s)
Ricin (R60)	72.0
Ricin A chain (RTA)	79.5
Ricin B chain (RTB)	81.0
MLI	61.5
MLI A chain (MLI-A)	79.5
MLI B chain (MLI-B)	69.0

are in close contact with each other although not overlapping. The $[AB]_2$ dimers are packed around a 6_5 screw axis in the unit cell. Self-association is supported by non-covalent interactions between two domain 1's of adjacent B chains as shown in Fig. 1. Domain 1 comprises residues 1–140 although residues 1–16 form an extended arm which covalently bonds to the A chain between Cys^5 and Cys^{247} respectively. Residues 17–140 form a β-trefoil structure which comprises a six stranded barrel and an assembly of three separate loops, α, β and γ (in the case of ricin described as a hairpin triplet), which caps the barrel [16]. The structural similarities between ricin and MLI B chains in this region reflect the sequence homology between the two proteins (Fig. 2). It is this triangular array of loops α, β and Δ, comprising residues 21–40 (α), 64–81 (β) and 107–124 (γ), that forms the dimer interface with the same triplet of loops from the adjacent molecule as shown in Fig. 3. Inspection of the interactions in the interface region indicates that contacts are formed predominantly between pairs of loops (α and γ′) and (γ and α′), with a less significant contribution from (β to β′). The strongest of these interactions appear to involve hydrophilic side chains in the α loop in the sequence Arg^{26}-Asp^{27}-Asp^{28}-Phe^{29}-Arg^{30}-Asp^{31}, with γ loop residues Thr^{116} and Thr^{117}. The β loop is stabilised by a disulphide bond at Cys^{64}–Cys^{81}. This disulphide bond is conserved in ricin, however the other internal disulphide bond of domain 1 at Cys^{20}–Cys^{39} of the α loop is not conserved in MLI where an evolutionary mutation has replaced Cys^{39} with serine. This mutation appears to allow the α loop more flexibility to take up its place at the dimer interface and suggests a reason why ricin does not form similar dimers. Domain 2 of MLI B and RTB have preserved the two internal disulphide bonds of the α and β loops.

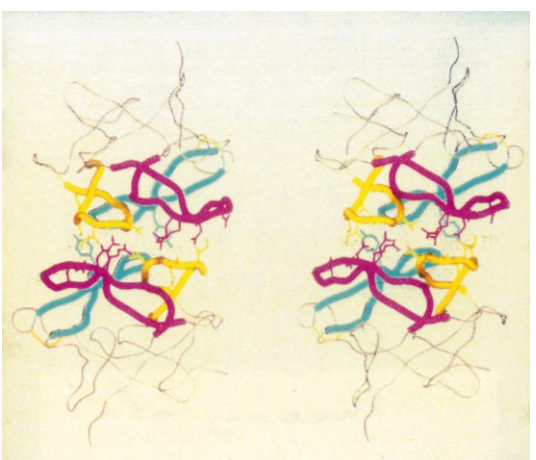

Fig. 3. Stereo view of MLI interface region between two B chain domain 1s. Loops forming the interacting triplets at the dimer interface are indicated (α = pink, residues 21–40; β = blue, residues 64–81; γ = yellow, residues 107–124). The disulphide bond on the β loop is indicated. The free Cys^{20} and Ser^{29} of MLI are indicated at the neck of the α loop. This is where a disulphide bridge is located in ricin. The γ loop is not stabilised by a disulphide bridge in either MLI or ricin. Interacting residues in the interface are depicted.

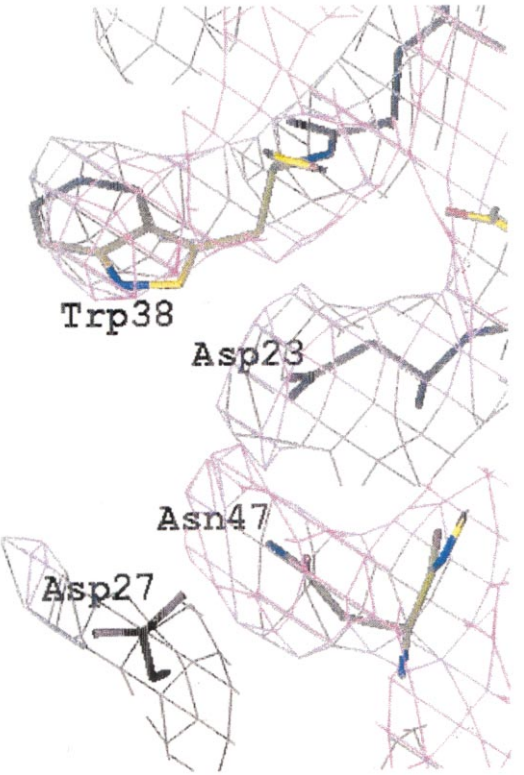

Fig. 4. $(2|F_o|-|F_c|)$ electron density map of MLI at the sugar binding site of domain 1.

Self-association of another type II RIP, abrin, has also been reported [17]. One carbohydrate binding site of the abrin B chain is occupied by a mannose group from the glycan chain of an adjacent molecule. According to the X-ray structure, there is no evidence that glycan chains in MLI play a similar role in dimerisation, although the sugar binding site of domain 1 is situated close to the dimer interface. The $(2|F_o|-|F_c|)$ electron density map indicates that this site is unoccupied (see Fig. 4). Four potential *N*-glycosylation sites were identified in the amino acid sequence of MLI. Position 112 of MLI A and 61, 96 and 136 of MLI B have Asn in the triplet Asn-X-Thr/Ser. Ricin, in contrast, has four occupied glycosylation sites, two on RTA at Asn^{10} and Asn^{236} [18,19] and two on RTB at Asn^{95} and Asn^{135} [12]. The Asn^{61}-Gly^{62}-Ser^{63} glycosylation site triplet in MLI B corresponds to a third glycosylated site at Asn^{357}-Cys^{358}-Ser^{359} in the homologue (RCA) ricin agglutinin [20]. RCA also forms an $[AB]_2$ homodimer. However, unlike MLI, the association is reinforced by a disulphide bridge between two adjacent A chains [21]. The functional role of dimerisation in RIP II toxins is as yet unclear; however, it has been suggested [22] that the ability of ricin, chemically modified with bound glycolipid ligands to self-associate, may mimic its behaviour at the cell surface during internalisation.

Acknowledgements: We thank the Royal Society for the award of a Joint Project Grant (JP875/AJM) and RFFI (97-04-48686). X-ray

data were collected at CLRC Synchrotron Laboratory, Daresbury, UK.

References

[1] Barbieri, L., Batteli, M.G. and Stirpe, F. (1993) Biochim. Biophys. Acta 1154, 237–282.
[2] Endo, Y., Chan, Y.-L., Lin, A., Tsurugi, K. and Wool, I.G. (1988) J. Biol. Chem. 263, 7917–7920.
[3] Lord, J.M. and Roberts, L.M. (1996) Plant Physiol. Biochem. 34, 253–261.
[4] Dietrich, J.G., Ribereau-Gayton, G., Jung, M.L., Franz, H., Beck, J.P. and Anton, R. (1992) Anticancer Drugs 3, 507–511.
[5] Pfueller, U. (1996) in: Grundlagen der Misteltherapie (Scheer, R., Becker, H. and Berg, P.A., Eds.), pp. 170–182, Hippokrates Verlag, Stuttgart.
[6] Hincha, D.K., Pfueller, U. and Schmitt, M.J. (1997) Planta 203, 140–144.
[7] Higgens, D.G., Thompson, J.D. and Gibson, T.J. (1996) Methods Enzymol. 226, 383–402.
[8] Agapov, I.I., Tonevitsky, A.G., Shamshiev, A.T., Pohl, P. and Palmer, R.A. (1997) FEBS Lett. 402, 91–93.
[9] Sweeney, E.C., Palmer, R.A. and Pfueller, U. (1993) J. Mol. Biol. 234, 1279–1281.
[10] Otwinowski, Z. (1993) Proceedings of the CCP4 Study Weekend (Sawyer, L., Isaacs, N. and Bailey, S., Eds.), pp. 56–62, SERC Daresbury Laboratory.
[11] CCP4 Protein Crystallography Programme Suite, CLRC, Daresbury Laboratory.
[12] Rutenber, E. and Robertus, J.D. (1991) Proteins 10, 260–269.
[13] Navaza, J. (1994) Acta Crystallogr. 50, 157–163.
[14] Wang, D., Driessen, H.P.C. and Tickle, I.J. (1991) J. Mol. Graph. 9, 50–52.
[15] Jones, T.A., Zou, Y.-J., Cowan, S.W. and Kjegaard, M. (1991) Acta Crystallogr. 47, 110–119.
[16] Murzin, A.G., Lesk, A.M. and Chothia, C. (1992) J. Mol. Biol. 223, 531–543.
[17] Tahirov, T.H., Lu, T.-H., Liaw, Y.-C., Chen, Y.-L. and Lin, J.-Y. (1994) J. Mol. Biol. 250, 354–367.
[18] Katzin, B.J., Collins, E.J. and Robertus, J.D. (1991) Proteins Struct. Funct. Genet. 10, 251–259.
[19] Kimura, Y., Kusuoku, H., Tada, M., Takagi, S. and Funatsu, G. (1990) Agric. Biol. Chem. 54, 157–162.
[20] Roberts, L., Lamb, F.I., Pappin, D.J.C. and Lord, J.M. (1985) Eur. J. Biochem. 148, 265–270.
[21] Sweeney, E.C., Tonevitsky, A.G., Temiakov, D.E., Agapov, I.I., Saward, S. and Palmer, R.A. (1997) Proteins Struct. Funct. Genet. 28, 586–589.
[22] Venkatesh, Y.P. and Lambert, J.M. (1997) Glycobiology 7, 329–335.