Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray diffraction analysis of HML, a lectin from the red marine alga *Hypnea musciformis*

Celso S. Nagano,[a]‡ Francisca Gallego del Sol,[a]‡ Benildo S. Cavada,[b] Kyria Santiago Do Nascimento,[b] Eudismar Vale Nunes,[c] Alexandre H. Sampaio[c] and Juan J. Calvete[a]*

[a]Instituto de Biomedicina de Valencia, CSIC, Valencia, Spain, [b]Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, CE 60451-970, Brazil, and [c]Laboratorio de Bioquímica Marinha, Departamento de Engenharia de Pesca, Universidade Federal do Ceará, Fortaleza, CE 60451-970, Brazil

‡ These authors contributed equally to this work and may both be considered 'first author'.

Correspondence e-mail: jcalvete@ibv.csic.es

Received 3 August 2004
Accepted 18 October 2005
Online 20 October 2005

HML, a lectin from the red marine alga *Hypnea musciformis*, defines a novel lectin family. Orthorhombic crystals of HML belonging to space group $P2_12_12_1$ grew within three weeks at 293 K using the hanging-drop vapour-diffusion method. A complete data set was collected at 2.4 Å resolution. HML is the first marine alga lectin to be crystallized.

1. Introduction

The recognition of carbohydrates by proteins underlies key cellular processes such as cell communication, host defence, fertilization, development, parasitic infection and tumour metastasis. Lectins are the carbohydrate-binding proteins of non-immune origin found in all types of living organisms that decipher the glycocodes encoded in the structure of glycans attached to soluble and integral cell-membrane glycoconjugates (Gabius & Gabius, 1997). Mechanisms for sugar recognition in microorganisms, plants and animals have evolved independently in diverse protein frameworks (Elgavish & Shaanan, 1997). The largest and best characterized lectin family is that from terrestrial plants and accumulating evidence indicates that the vast majority of these lectins can be classified into four large and three small families of structurally and evolutionary related proteins (Van Damme *et al.*, 1998; see also the 3D Lectin Database at http://webenligne.cermav.cnrs.fr/lectines/). The molecular structure and functional features of an increasing number of terrestrial plant lectins have been reported (Loris, 2002). However, and in marked contrast to higher land plant lectins, marine algal lectins have been isolated and characterized at a much lower pace since the first report of haemagglutinating activity in these organisms appeared more than 35 y ago (Boyd *et al.*, 1966). Moreover, to date biochemical and structural information on algal lectins is scarce and is available from only a few species, mainly owing to difficulties in their isolation and in obtaining sufficient material for study. Therefore, the functional and phylogenetic classification of these lectins remains obscure.

The available biochemical information indicates the existence of lectins belonging to different protein families in marine algae such as the green algae *Enteromorpha prolifera* (Ambrosio *et al.*, 2003) and *Ulva pertusa* (Wang *et al.*, 2004), the red marine algae *Bryothamnion triquetrum* (Calvete *et al.*, 2000), *Hypnea japonica* (Hori *et al.*, 2000), *H. musciformis* and *H. cervicornis* (Nagano *et al.*, 2005) and species of the genera *Eucheuma* (Kawakubo *et al.*, 1999) and *Ptilota* (Sampaio *et al.*, 1998, 1999, 2002). Moreover, the complete amino-acid sequences of only six algal lectins have been determined (BTL, Calvete *et al.*, 2000; HJA1 and HJA2, Hori *et al.*, 2000; UPL1, Wang *et al.*, 2004; HML and HCA, Nagano *et al.*, 2005) (Fig. 1).

HML (9357 ± 1 Da), a lectin isolated from the red marine alga *H. musciformis*, consists of a mixture of a 90-residue polypeptide containing seven intrachain disulfide bonds and two disulfide-bonded fragments generated by cleavage at the bond Arg50–Glu51 (Nagano *et al.*, 2005) (Fig. 1). The N-terminal (residues 1–47) and C-terminal (residues 48–90) fragments exhibit a large sequence similarity (Fig. 1) and HML can thus be regarded as a mixture of a single chain with internal homology and a two-chain protein of homologous polypeptides (Nagano *et al.*, 2005). HML exhibits neither discernible

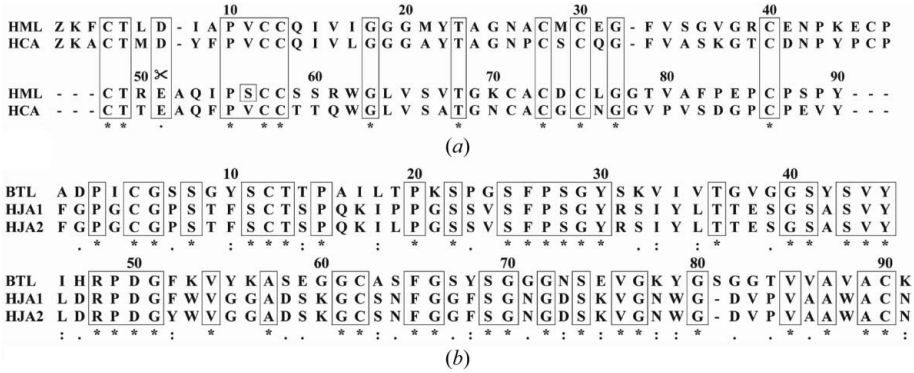

Figure 1
Amino-acid sequences of lectins belonging to two different protein families isolated from the taxonomically related species of red marine algae of the *Hypnea* and *Bryothamnion* genera. (*a*) HML and HCA, lectins from *H. musciformis* and *H. cervicornis*, respectively (Nagano *et al.*, 2005); (*b*) BTL, the lectin from *B. triquetrum* (Calvete *et al.*, 2000), and HJA1 and HJA2, isolectins A1 and A2 from *H. japonicus*, respectively (Hori *et al.*, 2000). Conserved residues are boxed. In (*a*), only those residues revealing internal domain duplication are boxed. The proteolytic cleavage at the Arg50–Glu51 peptide bond, which generates a two-chain lectin, is indicated by scissors. Z, pyroglutamic acid.

amino-acid sequence similarity with nor the cysteine-spacing pattern found in any other known protein structure, strongly indicating that HML belongs to a novel protein (lectin) family (Nagano *et al.*, 2005). The high cysteine (disulfide bond) content of HML is an unusual feature for lectin structures. Of the plant lectins, the only other known example of cysteine-rich proteins are the chitin-binding lectins, which are made up of hevein domains (Van Damme *et al.*, 1998). Here, we report the first crystallization and the preliminary X-ray diffraction analysis from a marine alga lectin.

2. Materials and methods

2.1. Purification of *H. musciformis* lectin (HML)

Specimens of the red alga *H. musciformis* were collected on Pacheco beach on the Atlantic coast of the Ceará State of Brazil.

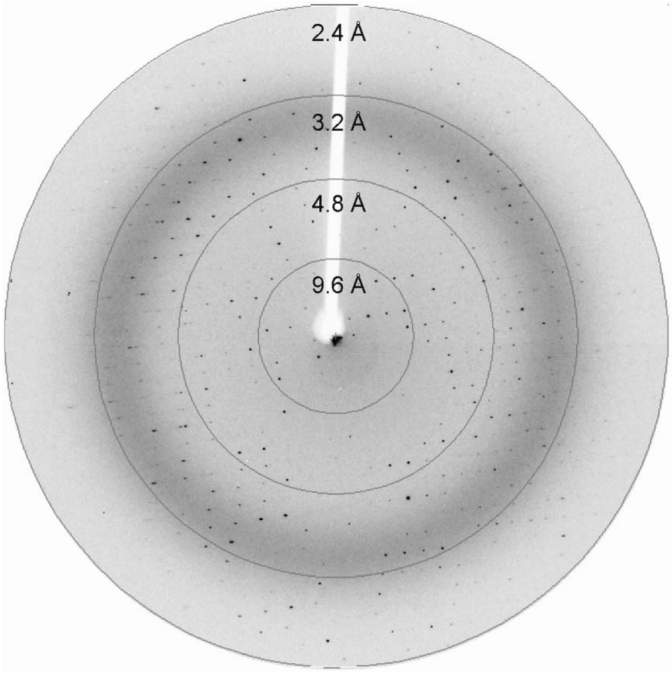

Figure 2
Diffraction pattern of an orthorhombic crystal of HML at 293 K showing maximum resolution at 2.4 Å.

HML was purified as described by Nagano *et al.* (2005) and its purity was assessed by SDS–PAGE and MALDI–TOF mass spectrometry [using an Applied Biosystems Voyager DE-PRO instrument operating at 25 kV accelerating voltage in the linear mode, with 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid) saturated in 70% acetonitrile and 0.1% TFA used as the matrix].

2.2. Analytical ultracentrifugation

The apparent molecular mass of the HML in solutions of different pH was determined by analytical ultracentrifugation at 293 K in a Beckman XL-A centrifuge with UV absorption scanner optics using an AN-50 Ti eight-hole rotor and charcoal-filled EPON six-channel centrepieces. This setup allows the simultaneous analysis of five different samples. The lectin was dissolved at 1.0 mg ml^{-1} in 20 m*M* NH_4HCO_3 pH 8.5, 20 m*M* HEPES pH 7.5, 20 m*M* MES pH 6.5 or 5.5 and 20 m*M* sodium acetate pH 4.5. Molar masses were determined by sedimentation-diffusion equilibrium experiments using short (approximately 3 mm) sedimentation columns. To avoid differences in apparent molecular masses arising from rotor speed-dependent weighting, all experiments were carried out at the same speed (15 000 rev min^{-1}). When the measured concentration profile remained unchanged for at least 12 h, equilibrium was assumed to be attained. The equilibrium concentration gradient for a single species is described by

$$\ln c(r) - \ln c(m) = [M(1 - \overline{\upsilon}\rho)/2RT]\omega^2(r^2 - m^2),$$

where $c(r)$ and $c(m)$ are the concentrations at radius r and at the meniscus (radius m), respectively, M is the molar mass of the solute, $\overline{\upsilon}$ is the partial specific volume of the solute (assumed to be 7.35×10^{-2} ml g^{-1}), ω is the angular speed of the rotor (rad s^{-1}) and R and T are the gas constant (8.314 J mol^{-1} K^{-1}) and the temperature (in K), respectively. Apparent molar masses were determined by fitting this function to the measured radial distribution of the concentration gradient at equilibrium using the program *EQASSOC* (Minton, 1994) provided by the manufacturer. Blank buffer absorption was determined after overspeeding to sediment all material to the bottom of the cell.

2.3. Crystallization, data collection and processing

The sparse-matrix method (Jancarik & Kim, 1991) using Crystal Screens I and II supplied by Hampton Research (California, USA) was utilized to perform initial screening of the crystallization condi-

Table 1
Data-collection statistics.

Data in parentheses are for the highest resolution shell (2.52–2.40 Å).

Space group	$P2_12_12_1$
Unit-cell parameters	
a (Å)	44.89
b (Å)	64.64
c (Å)	104.90
α (°)	90
β (°)	90
γ (°)	90
Resolution range (Å)	50–2.4
Total reflections	108681
Unique reflections	10746
Data completeness (%)	87.5 (87.5)
$R_{\rm merge}$† (%)	11.1 (33.7)
Multiplicity	5.4 (5.1)
$I/\sigma(I)$	5 (1.7)

† $R_{\rm merge} = \sum|I_{\rm obs} - \langle I\rangle|/\sum\langle I\rangle$.

tions. Crystals were grown at 295 K by the vapour-diffusion method using 2 µl hanging drops composed of equal volumes of protein solution (15 mg ml^{-1} in water) and reservoir buffer (10% PEG 8000, 8% ethylene glycol, 100 m*M* HEPES pH 7.5). Crystals suitable for diffraction experiments (maximum dimensions 0.6 × 0.4 × 0.2 mm) grew within three weeks. For X-ray diffraction, crystals were flash-frozen at 100 K in a nitrogen-gas stream. Reservoir buffer containing 20% ethylene glycol proved to be a suitable cryoprotectant. The crystals were transferred directly from the drop to the cryoprotectant solution and were allowed to equilibrate for approximately 10 s. Single crystals were then mounted in nylon loops and rapidly transferred to the cryostream. The cryoprotected crystals were analysed using synchrotron radiation on BM16 at the ESRF (Grenoble). A complete data set was collected using a CCD (MAR Research) in 100 frames with an oscillation range of 1°. The data were indexed, integrated and scaled using *MOSFLM* (Leslie, 1997) and *SCALA* (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

The average flotation masses $[M(1 - \overline{\upsilon}\rho)]$ of HML determined in buffers of different pH in the range 4.5–8.5 by sedimentation-equilibrium analytical centrifugation were 3170–3400. These values correspond to apparent molecular weights of 12–13 kDa, which when compared with the MALDI–TOF mass of HML, 9357, indicate that the major molecular species is monomeric. The sedimentation curves at equilibrium could be adjusted to a single species at the acidic pH values [correlation coefficients 0.8 (pH 4.5), 0.97 (pH 5.5) and 1.0 (pH 6.5)], whereas formation of higher order species was detected in solutions of basic pH. The tendency to self-associate into higher order aggregates is a common phenomenon among lectins from diverse families.

Crystals of the *H. musciformis* lectin were grown at pH 7.5 in hanging drops by the vapour-diffusion method using a mixture of 10% PEG 8000 and 8% ethylene glycol as precipitant. The crystals belong to the orthorhombic space group $P2_12_12_1$ (unit-cell parameters a = 44.89, b = 64.64, c = 104.90 Å) and diffracted to a maximum resolution of 2.4 Å (Fig. 2). Data in the 55.0–2.4 Å resolution range were scaled using *SCALA* to an $R_{\rm merge}$ of 0.11 (0.33 in the highest resolution shell) with an $I/\sigma(I)$ of 5 (1.7 in the highest resolution shell) (Table 1). Considerations of the packing density support the presence of four molecules in the asymmetric unit, which would correspond to a Matthews coefficient ($V_{\rm M}$) of 2.0 Å^3 Da^{-1} and a solvent content of 39.1%.

The crystal structure of HML could not be solved using a molecular-replacement approach owing to the lack of a search model. HML binds GalNAc/Gal substituted with a neutral sugar through 1–3, 1–4 or 1–2 linkages in *O*-linked mucin-type glycans and Fuc(α1–6)GlcNAc of N-linked glycoproteins (Nagano *et al.*, 2005). Analysis of suitable heavy-atom derivatives and the possibility of MAD phasing using either a sugar derivative as a carrier of an anomalous X-ray scatterer (Buts *et al.*, 2003; Gallego del Sol *et al.*, 2005) or the anomalous signal of sulfur (Ramagopal *et al.*, 2003; Usón *et al.*, 2003) are under way.

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Financiadora de Estudos e Projetos (FINEP), Fundação Cearense de Amparo à Pesquisa (FUNCAP) and grant BFU2004-01432/BMC from the Ministerio de Educación y Ciencia, Madrid (Spain). CSN is the recipient of a fellowship from the Coordenação Aperfeiçoamento de Pessoal de Nivel Superior (CAPES). AHS and BSC are senior investigators of CNPq.

References

Ambrosio, A. L., Sanz, L., Sánchez, E. I., Wolfenstein-Todel, C. & Calvete, J. J. (2003). *Arch. Biochem. Biophys.* **415**, 245–250.
Boyd, W. C., Almodovar, L. R. & Boyd, L. G. (1966). *Transfusion*, **6**, 82–83.
Buts, L., Loris, R., De Genst, E., Oscarson, S., Lahmann, M., Brosens, E., Wyns, L., De Greve, H. & Bouckaert, J. (2003). *Acta Cryst.* D**59**, 1012–1015.
Calvete, J. J., Costa, F. H. F., Saker-Sampaio, S., Moreno-Murciano, M. P., Nagano, C. S., Cavada, B. S., Grangeiro, T. B., Ramos, M. V., Bloch, C. Jr, Silveira, S. B., Freitas, B. T. & Sampaio, A. H. (2000). *Cell. Mol. Life Sci.* **57**, 343–350.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Elgavish, S. & Shaanan, B. (1997). *Trends Biochem. Sci.* **22**, 462–467.
Gabius, H.-J. & Gabius, S. (1997). Editors. *Glycoscience: Status and Perspectives.* Weinheim: Chapman & Hall.
Gallego del Sol, F., Gómez, J., Hoos, S., Nagano, C. S., Cavada, B. S., England, P. & Calvete, J. J. (2005). *Acta Cryst.* F**61**, 326–331.
Hori, K., Matsubara, K. & Miyazawa, K. (2000). *Biochim. Biophys. Acta*, **1474**, 226–236.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kawakubo, A., Makino, H., Ohnishi, J.-I., Hirohara, H. & Hori, K. (1999). *J. Appl. Phycol.* **11**, 149–156.
Leslie, A. G. W. (1997). *MOSFLM User's Guide: MOSFLM Version* 5.50. Laboratory of Molecular Biology, Cambridge.
Loris, R. (2002). *Biochim. Biophys. Acta*, **1572**, 198–208.
Minton, A. P. (1994). *Modern Analytical Ultracentrifugation*, edited by T. M. Schuster & T. M. Laue, pp. 81–93. Boston, MA, USA: Birkhauser.
Nagano, C. S., Debray, H., Nascimento, K. S., Pinto, V. P. T., Cavada, B. S., Saker-Sampaio, S., Farias, W. R. L., Sampaio, A. H. & Calvete, J. J. (2005). *Protein Sci.* **14**, 2167–2176.
Ramagopal, U. A., Dauter, M. & Dauter, Z. (2003). *Acta Cryst.* D**59**, 1020–1027.
Sampaio, A. H., Rogers, D. J. & Barwell, C. J. (1998). *Phytochemistry*, **48**, 765–769.
Sampaio, A. H., Rogers, D. J., Barwell, C. J., Saker-Sampaio, S., Costa, F. H. F. & Ramos, M. V. (1999). *J. Appl. Phycol.* **10**, 539–546.
Sampaio, A. H., Rogers, D. J., Barwell, C. J., Saker-Sampaio, S., Nascimento, K. S., Nagano, C. S. & Farias, W. R. L. (2002). *J. Appl. Phycol.* **14**, 489–495.
Usón, I., Schmidt, B., von Bulow, R., Grimme, S., von Figura, K., Dauter, M., Rajashankar, K. R., Dauter, Z. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 57–66.
Van Damme, E. J. M., Peumans, W. J., Barre, A. & Rougé, P. (1998). *Crit. Rev. Plant Sci.* **17**, 575–692.
Wang, S., Zhong, F.-D, Zhang, Y.-J., Wu, Z.-J., Lin, Q.-Y. & Xie, L. H. (2004). *Acta Biochim. Biophys. Sin.* **36**, 111–117.