Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of full-length *Toxascaris leonina* galectin with two carbohydrate-recognition domains

Mi Suk Jeong,[a] Hyun Gi Hwang,[a] Hak Sun Yu[b] and Se Bok Jang[a]*

[a]Department of Molecular Biology, College of Natural Sciences, Pusan National University, Jangjeon-dong, Geumjeong-gu, Busan 609-735, Republic of Korea, and [b]Department of Parasitology, School of Medicine, Pusan National University, Beomeo-ri, Mulgeum-eup, Yangsan-si, Gyeongsangnam-do 626-870, Republic of Korea

Correspondence e-mail: sbjang@pusan.ac.kr

Received 13 July 2012
Accepted 31 October 2012

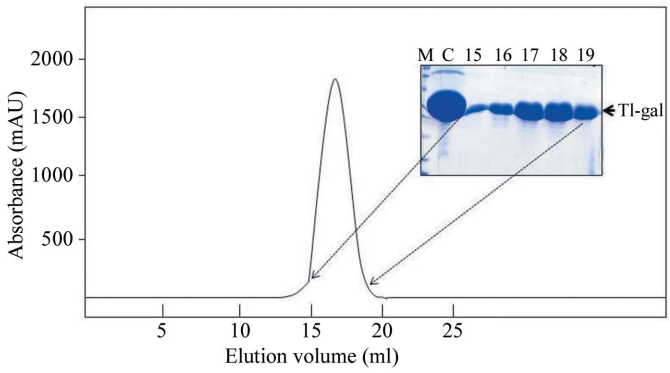
PDB Reference: Tl-galectin, 4hl0

The full-length crystal structure of *Toxascaris leonine* galectin (Tl-gal), a galectin-9 homologue protein, was determined at a resolution of 2.0 Å. Galectin-9 exhibits a variety of biological functions, including cell aggregation, eosinophil chemoattraction, activation and apoptosis of murine thymocytes, T cells and human melanoma cells. Similar to this galectin, Tl-gal may function as a regulatory molecule in the host immune system; however, no molecular or structural information has been reported for Tl-gal. Moreover, until now, there have been no reports of a full-length galectin structure. There are two molecules of Tl-gal per asymmetric unit in space group $P2_12_12_1$, and the N-terminal and C-terminal carbohydrate-recognition domains (NCRD and CCRD) of Tl-gal are composed of six-stranded β-sheets and five-stranded β-sheets with a short α-helix. The NCRD of Tl-gal resembles that of human galectin-7 and its CCRD resembles human galectin-9, but the residues in the interface and loop regions of the NCRD and CCRD are flexible and are related to interaction. Engagement of the T-cell immunoglobulin mucin-3 (Tim-3) immunoglobulin variable (IgV) domain by a galectin-9 ligand is known to be important for appropriate termination of T-helper 1 immune responses. To investigate the binding site of Tl-gal, the interaction between Tl-gal and Tim-3 was modelled. Tim-3 is docked into a major groove of the Tl-gal structure, which is larger and deeper than the minor groove. The structural information presented here will provide insight into the development of novel anti-inflammatory agents or selective modulators of immune response.

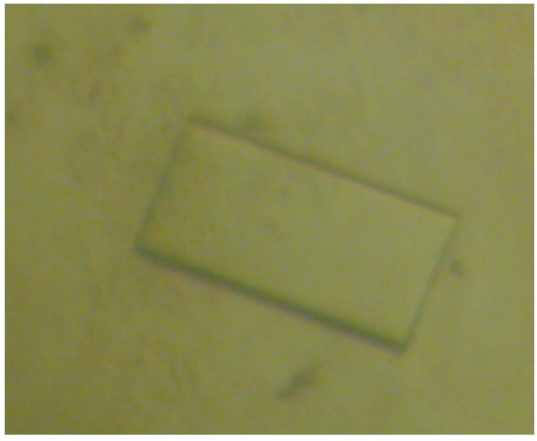
1. Introduction

Galectins cross-link glycoproteins to form dynamic microdomains or lattices that regulate various mediators of cell adhesion, migration, proliferation, survival and differentiation (Boscher *et al.*, 2011). It has been proposed that galectins play various biological roles throughout the organism that are related to the regulation of immunity and inflammation, the progression of cancer and specific developmental processes (Rabinovich *et al.*, 2002; Bidon-Wagner & Le Pennec, 2004; Hughes, 2004). Some galectins have recently come into focus in the reproductive sciences and perinatal medicine because they are highly expressed at the maternal–foetal interface (Than *et al.*, 2012). Indeed, the functional significance of galectins in eutherian pregnancies has been documented and their dysregulated expression is observed in the 'great obstetrical syndromes' (Blois *et al.*, 2007; Than *et al.*, 2009).

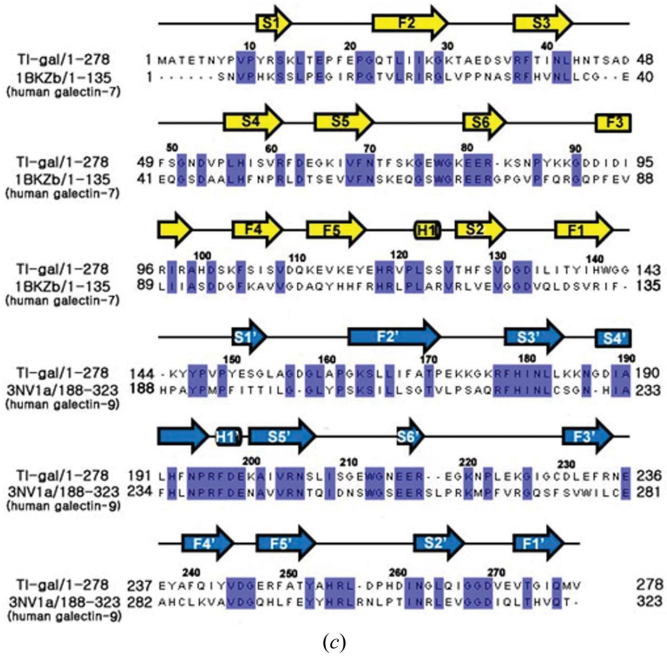
Galectin-9 was originally isolated from tumour cells in Hodgkin's disease, a condition characterized by blood and tissue eosinophilia (Türeci *et al.*, 1997). The expression of galectin-9, which is encoded by three genes in humans, is predominant in immune cells, intestinum and endometrium/



(*a*)



(*b*)



(*c*)

Figure 1
Elution, crystal and sequence alignment of Tl-gal. (*a*) Elution profile of Tl-gal from a Superdex 200 10/300 GL column. The single monodisperse peak at 16.5 ml is consistent with the predicted Tl-gal. (*b*) Crystal of Tl-gal. (*c*) Sequence alignments of the NCRD and CCRD of Tl-gal with those of human galectins (galectin-7, 1BKZb; galectin-9, 3NV1a) showing the secondary structure. Secondary-structure elements are shown as arrows (β-sheets) and a rectangle (α-helix). Loops are shown as black lines. Residues that are conserved in the two species are shown in blue. Every tenth residue is indicated by a number.

deciduas (Su *et al.*, 2004). The chemoattractant activity of galectin-9 depends on its carbohydrate-binding activity and requires the N- and C-terminal carbohydrate-recognition domains (CRDs; Matsushita *et al.*, 2000). The sequence similarity between the N- and C-terminal CRDs is 35%. Galectin-9 has also been identified as a T-cell immunoglobulin mucin 3 (Tim-3) ligand that specifically recognizes a carbohydrate motif on the Tim-3 immunoglobulin variable domain (Zhu *et al.*, 2005; Cao *et al.*, 2007). Furthermore, the *in vivo* administration of recombinant galectin-9 during an immune response results in selective deletion of IFN-γ-producing cells and the amelioration of autoimmune disease.

A gene encoding a galectin-9 homologue (Tl-gal) was previously isolated from an adult worm of the canine gastrointestinal nematode parasite *Toxascaris leonine via* random cDNA library sequencing (Kim *et al.*, 2010). Similar to the galectin, Tl-gal may function as a regulatory molecule in the host immune system; however, molecular and structural information has not yet been elucidated for Tl-gal. Additionally, there have been no reports of a full-length galectin structure to date. Here, we report the first crystal structure of full-length Tl-gal and discuss the relationship between the structure and function of Tl-gal. This information will provide insight into the design of novel regulatory molecules for the treatment of inflammatory, allergic and immune diseases.

2. Materials and methods

2.1. Protein expression and purification

The full-length Tl-gal gene (amino acids 1–278) was incorporated into plasmid pET-28a (Novagen) in *Escherichia coli* BL21 (DE3) cells, grown in LB (Luria–Bertani) medium and induced with 0.5 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) for 16 h at 293 K. The cells were harvested by centrifugation and the cell pellets were resuspended in lysis buffer *A* (50 m*M* Tris–HCl pH 7.5, 200 m*M* NaCl). The cells were then disrupted by sonication on ice, after which the soluble Tl-gal supernatant was loaded onto an Ni–NTA column (Amersham Pharmacia Biotech) that had been pre-equilibrated with buffer *A*. The column was washed with buffer *A* containing 20 m*M* imidazole, after which the bound protein was eluted using buffer *A* containing 200 m*M* imidazole. Purified fractions of the Tl-gal protein from the Ni–NTA column were then mixed and purified by gel-filtration chromatography. Gel filtration was performed by FPLC using a Superdex 200 10/300 GL column (Amersham Pharmacia Biotech) equilibrated in buffer *A*, after which the fractions were collected (Fig. 1*a*). All purification steps were assessed by 12% SDS–PAGE and Coomassie Blue staining. The purified Tl-gal protein was confirmed to run as a single band with a molecular weight of 32 kDa on SDS–PAGE.

2.2. Crystallization

Crystallization was performed by the hanging-drop vapour-diffusion method using a crystallization screening kit at 293 K. Briefly, a hanging drop was prepared by mixing equal volumes

Table 1
Crystallographic statistics.

Values in parentheses are for the highest resolution shell.

Space group	$P2_12_12_1$
Unit-cell parameters (Å, °)	a = 78.6, b = 82.8, c = 123.0, $\alpha = \beta = \gamma = 90$
Resolution (Å)	50.0–2.0 (2.03–2.00)
Completeness (%)	95.0 (99.3)
Observed reflections	97056
Unique reflections	54560
$\langle I/\sigma(I)\rangle$	25.7 (3.9)
$R_{\rm merge}$† (%)	10.2 (37.2)
Multiplicity	9.3
$R_{\rm cryst}/R_{\rm free}$‡ (%)	25.4/29.3
No. of protein/water atoms	556/443
R.m.s.d. bond lengths (Å)/angles (°)	0.007/1.5
Average B factor (Å^2)	58.4
Ramachandran plot, residues in (%)	
Most favoured region	85.2
Additional allowed region	11.9
Generously allowed region	1.3

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the measured reflection and $\langle I(hkl)\rangle$ is the mean value of all symmetry-related reflections. ‡ $R_{\rm cryst} = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$, where $F_{\rm obs}$ and $F_{\rm calc}$ denote the observed and the calculated structure-factor amplitudes, respectively. $R_{\rm free} = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$ for a test data set of about 5% of the total reflections that were randomly chosen and set aside prior to refinement.

(1.0 µl each) of the protein solution and the reservoir solution. The crystals were grown using a reservoir consisting of 10% 2-propanol, 0.1 *M* MES buffer pH 6.0, 0.2 *M* calcium acetate. Under these conditions, rectangular crystals appeared after about 3 d and grew as single crystals with typical dimensions of 0.2 × 0.1 × 0.02 mm within two weeks (Fig. 1*b*).

2.3. Data collection and structure determination

Data for full-length Tl-gal were collected on beamline 7A of the Pohang Light Source (PLS), Republic of Korea. Before data collection, the crystal was briefly soaked in a cryoprotectant solution consisting of the precipitant solution with 20% glycerol. The crystal was flash-cooled in a liquid-nitrogen stream at 100 K. Diffraction data were then obtained and were processed using the *HKL*-2000 package (Otwinowski & Minor, 1997; Table 1).

The structure of Tl-gal was determined by the molecular-replacement method using the programs *AMoRe* and *EPMR* (Navaza, 1994; Kissinger *et al.*, 1999). The C-terminal domain (amino acids 145–276) of human galectin-9 (amino acids 188–323; PDB entry 3nv1; Yoshida *et al.*, 2010), with a sequence identity of 40%, was used as a starting search model. The search was carried out using data obtained between 12.0 and 4.5 Å resolution. The first molecule had a correlation coefficient of 11.7 (R = 53.6%) and solution of the second molecule resulted in a correlation coefficient of 19.6 (R = 52.5%) using *AMoRe*. For determination of the structure of the N-terminal domain of Tl-gal, the N-terminal domain of human galectin-7 (amino acids 1–135; PDB entry 1bkz; Leonidas *et al.*, 1998), with a sequence identity of 32%, was used as a search model in *EPMR*. The positioning of the N-terminal domain was determined by a six-parameter search using data with resolution between 12.0 and 4.5 Å. The positioning of this domain of the first molecule had a correlation coefficient of 31.7 (R = 82.7%) and solution of the second molecule resulted in a correlation coefficient of 53.6 (R = 69.4%).

The model was improved by iterative model building using the program *O* (Jones *et al.*, 1991). As shown in Table 1, final refinement to an R and $R_{\rm free}$ of 25.4% and 29.4%, respectively, was accomplished using the *CNS* package (Brünger *et al.*, 1998). The model exhibited good geometry, with r.m.s. (root-mean-square) deviations from ideal bond lengths and angles of 0.007 Å and 1.503°, respectively. The Ramachandran plot indicated that all amino-acid residues were in the generously allowed regions apart from four residues. The structural representations in Figs. 2–4 were generated using *MolScript*, *GRASP*, *Raster3D* and *PyMOL* (Kraulis, 1991; Nichols & Yanofsky, 1979; Merritt & Murphy, 1994; DeLano, 2002).

3. Results and discussion

3.1. Overall structure of full-length Tl-gal

We have successfully determined the first full-length crystal structure of *T. leonine* galectin (Tl-gal; amino acids 1–278) at a resolution of 2.0 Å (Fig. 2*b*). There were two molecules of Tl-gal per asymmetric unit in space group $P2_12_12_1$, with a solvent content of 59.1% (Fig. 2*c*). The galectin subfamilies have been proposed to be proto, chimera and tandem-repeat types based on their domain organizations (Hirabayashi & Kasai, 1993). Tl-gal resembles other tandem-repeat-type galectins, containing two carbohydrate-recognition domains in the N- and C-terminal regions (Fig. 2*a*). A linker region (residues 142–150) comprising a long loop is present between the N-terminal CRD (NCRD) and the C-terminal CRD (CCRD). The crystal structure shows that the NCRD and CCRD of Tl-gal are composed of six-stranded β-sheets (S1–S6 and S1′–S6′) and five-stranded β-sheets (F1–F5 and F1′–F5′) with a short α-helix (Fig. 2*d*). The Tl-galectin jelly-roll fold in each CRD is composed of two antiparallel β-sheets, which together form a β-sandwich arrangement. The short α-helix (H1) in the NCRD of Tl-gal is located between F5 and S2 as in the structure of hG9N, but interestingly H1′ in the Tl-gal CCRD is located between S4′ and S5′, thus differing from hG9C (Yoshida *et al.*, 2010).

3.2. Interdomain communication between the NCRD and the CCRD

The residues at the interface of the NCRD and CCRD may play an important role in interdomain communication. Residues 96–112 and 226–230 are related to the interaction and may stabilize the two CRD structures (Fig. 2*e*). The distances between the O atoms of Glu112, Gln110 and Arg96 in the NCRD and the main-chain amide groups of residues Gly226, Gly228, Cys229 and Asp230 range from 2.6 to 3.2 Å. A salt bridge is formed by Arg96 and Asp230, and this ionic interaction stabilizes the two CRDs. Moreover, the N atom of His140 maintains a stable geometry with Gln276 *via* a hydrogen bond. Residues 142–150 are composed of a flexible linker, and the N- and C-terminal regions are in close contact

to stabilize their charges (Fig. 3*a*). The electron density was of excellent quality, but residues 1–6 and 45–52 in the two molecules of the $P2_12_12_1$ structure had poor electron density. We confirmed the density maps using the other crystal structure of Tl-gal in space group $P1$ at 2.5 Å resolution (Figs. 3*b* and 3*c*). The electron density observed in the $P1$ structure was clearly visible in the regions of residues 1–6 and 45–52 of the $P2_12_12_1$ structure, but the R values for the $P1$ crystal structure with three molecules per asymmetric unit were slightly higher and the resolution was lower than that of the $P2_12_12_1$ crystal structure. These residues may be considered to be very flexible regions. For the crystal structures of Tl-gal in space groups $P2_12_12_1$ and $P1$, the r.m.s. deviation of C^{α} atoms between the two forms was 0.5–0.7 Å. Overall, these comparisons suggest that the two structures are very similar.

The structure of full-length Tl-gal had two hydrophobic cores in the NCRD and CCRD (Figs. 4*a* and 4*b*). Phe, Ile and Leu in the hydrophobic core are structurally important residues that are situated in a completely hydrophobic environment in the interior of the protein. The β-strands form an intermolecular hydrophobic core together with the opposite β-strands. In the structure, the hydrophobic core is surprisingly densely packed with side chains in the interior of the protein. The native conformation of a protein is maintained by numerous noncovalent interactions resulting from hydrogen bonds, salt bridges and hydrophobic effects. Hydrophobic effects have been shown to be a result of hydrophobic residues in the interior of a protein and to contribute to its conformational stability (Kellis *et al.*, 1988; Shortle *et al.*, 1990).

3.3. Structural comparison with other galectins

The sequence of the Tl-gal protein showed 35% homology to the human galectin-9 protein. However, the full-length

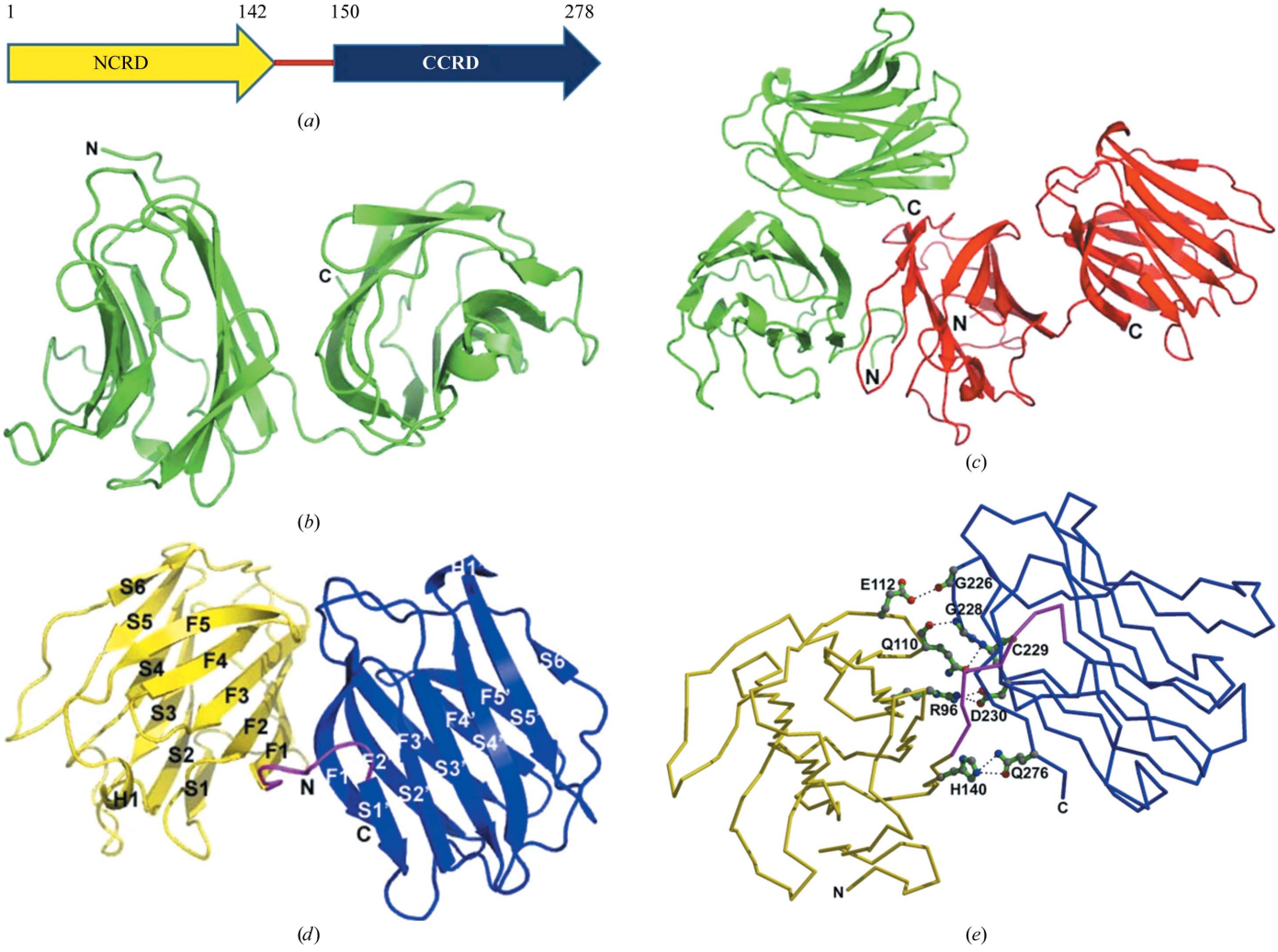


Figure 2
Structure and schematic diagram of full-length Tl-gal. (*a*) Linear representation of Tl-gal showing the boundaries of the NCRD, connection region and CCRD. The sequences of the NCRD, connection region and CCRD are residues 1–142, 142–150 and 150–278, respectively. (*b*) The NCRD and CCRD structures of Tl-gal are shown as two different shaped holes that are surrounded by β-sheets. (*c*) Crystal packing between one molecule (green) and the other molecule (red). (*d*) Ribbon diagram of a Tl-gal monomer. The NCRD, connection region and CCRD are shown in yellow, magenta and blue, respectively. (*e*) C^{α} trace of Tl-gal showing the boundaries of the NCRD (yellow), connection region (magenta) and CCRD (blue). The interface regions between the NCRD and CCRD are shown with the interacting residues in ball-and-stick representation.

human galectin-9 structure has not yet been determined. Sequence alignments of the NCRD and CCRD of Tl-gal with those of human galectins are shown in Fig. 1(*c*). The NCRD of Tl-gal (residues 7–141) is approximately 32% homologous to that of human galectin-7 (amino acids 1–135) and the r.m.s. deviation of the Tl-gal NCRD and the human galectin-7 NCRD was 4.22 Å (Fig. 4*c*). Superposition of the two structures revealed that residues with r.m.s. deviations exceeding 5.0 Å are located in the loop and strands S4 and S5 (residues 48–62 and 65–82). These results indicate that large conformational changes are present in the connection-loop region exposed to solvent between the β-strands.

The Tl-gal CCRD (residues 150–278) is homologous to the human galectin-9 CCRD (residues 189–321), with an amino-acid sequence similarity of 40%. The r.m.s. deviation of the Tl-gal CCRD and the human galectin-9 CCRD was 3.88 Å (Fig. 4*c*). Superposition of the two structures revealed that residues with r.m.s. deviations that exceed 4.0 Å are located in the loop and strands S5′ and F4′ (residues 202–209 and 239–242). The r.m.s. deviations of residues 257 and 277 in the loop and strands S2′ and F1′ are about 5–7 Å. A superposition of the C^{α} trace of the Tl-gal structure on those of the human galectin-7 and galectin-9 structures is shown in Fig. 4(*d*). The overall r.m.s. deviation of the Tl-gal and human galectin-7 NCRDs was larger than that of the Tl-gal and human galectin-9 CCRDs. The β-strand regions showed relatively small conformational changes. The numbers of C^{α} atoms aligned in the superposition of the C^{α} traces of the Tl-gal on the human galectin-7 and galectin-9 structures were 135 and 133, respectively.

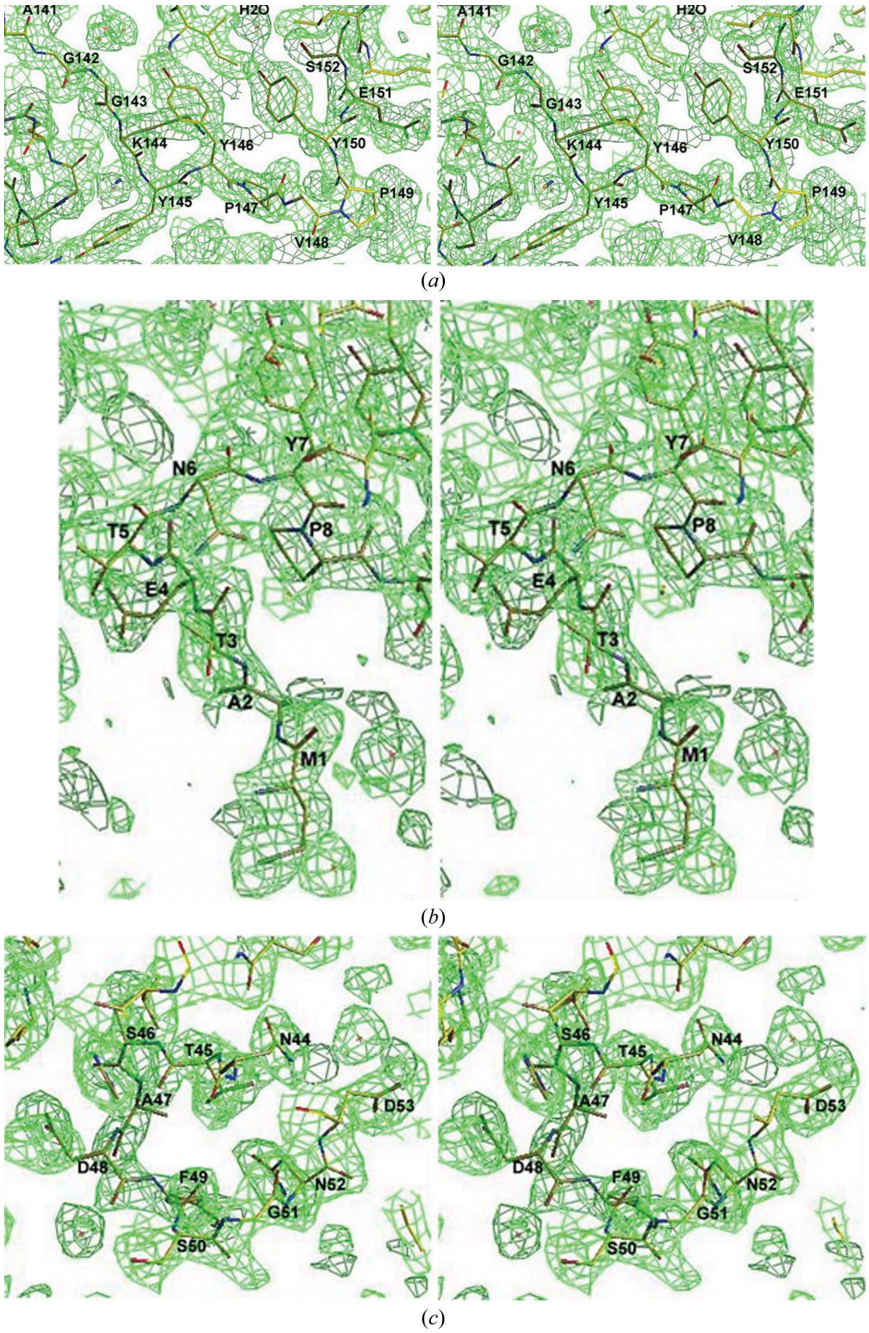


Figure 3
Electron-density maps of Tl-gal. (*a*) An electron-density map covering the linker between the NCRD and CCRD: stereoview of a $2F_o - F_c$ electron-density map of the residues contoured at 1σ. (*b*) The electron-density map of residues 1–6 in space group *P*1. (*c*) The electron-density map of residues 45–52 in space group *P*1.

3.4. Model of intermolecular communication between Tl-galectin and Tim-3

Engagement of the T-cell immunoglobulin mucin-3 (Tim-3) immunoglobulin variable (IgV) domain by galectin-9 is known to be important for appropriate termination of T-helper 1 immune responses (Zhu *et al.*, 2005). To investigate the binding site of the Tl-gal protein, we modelled the interaction between Tl-gal and Tim-3 using the known structure of Tim-3 (amino acids 1–109; PDB entry 2oyp; Cao *et al.*, 2007). Most positive charges in the Tim-3 structure were densely packed in the binding pocket. The negatively charged Tl-gal binding pocket binds to the positively charged pocket of Tim-3 (Fig. 5*a*). Because the Tl-gal protein surface features dictate its mode of interaction with partners, the electrostatic surface differential is critical for interactions with partner proteins. Tl-gal possesses distinctive major and minor grooves. Tim-3 is docked into a major groove of the Tl-gal structure (Fig. 5*b*). The position for interaction of Tim-3 and Tl-gal is predicted to be similar to the position for intercommunication of Tl-gal in the two molecules (Fig. 2*c*). The major groove of Tl-gal is larger and deeper than the minor groove and accommodates the side chains and main chains of loop regions of the Tim-3 molecule. Tl-gal

has hydrophobic residues and its structure contains distinctive NCRD and CCRD hydrophobic holes. Most of the hydrophobic residues are located in the two hydrophobic holes. We suggest that the long patch between the β-strands is the major contact surface for binding of a partner protein. The minor groove of Tl-gal is fully occupied by water molecules. The Tl-gal structure is composed of many β-strands, which are connected by loops, most of which are on the surface of the molecule. Tl-gal contains many solvent-accessible polar and charged residues. Most of the polar residues (26%) and charged residues (29%), including negatively charged (Asp and Glu) and positively charged (Lys and Arg) residues, are solvent-exposed. The loop regions exposed to the solvent are rich in charged and polar hydrophilic residues. The function of the loop regions is to participate in the formation of binding sites and protein active sites. The loop regions that connect two adjacent antiparallel β-strands are hairpin loops.

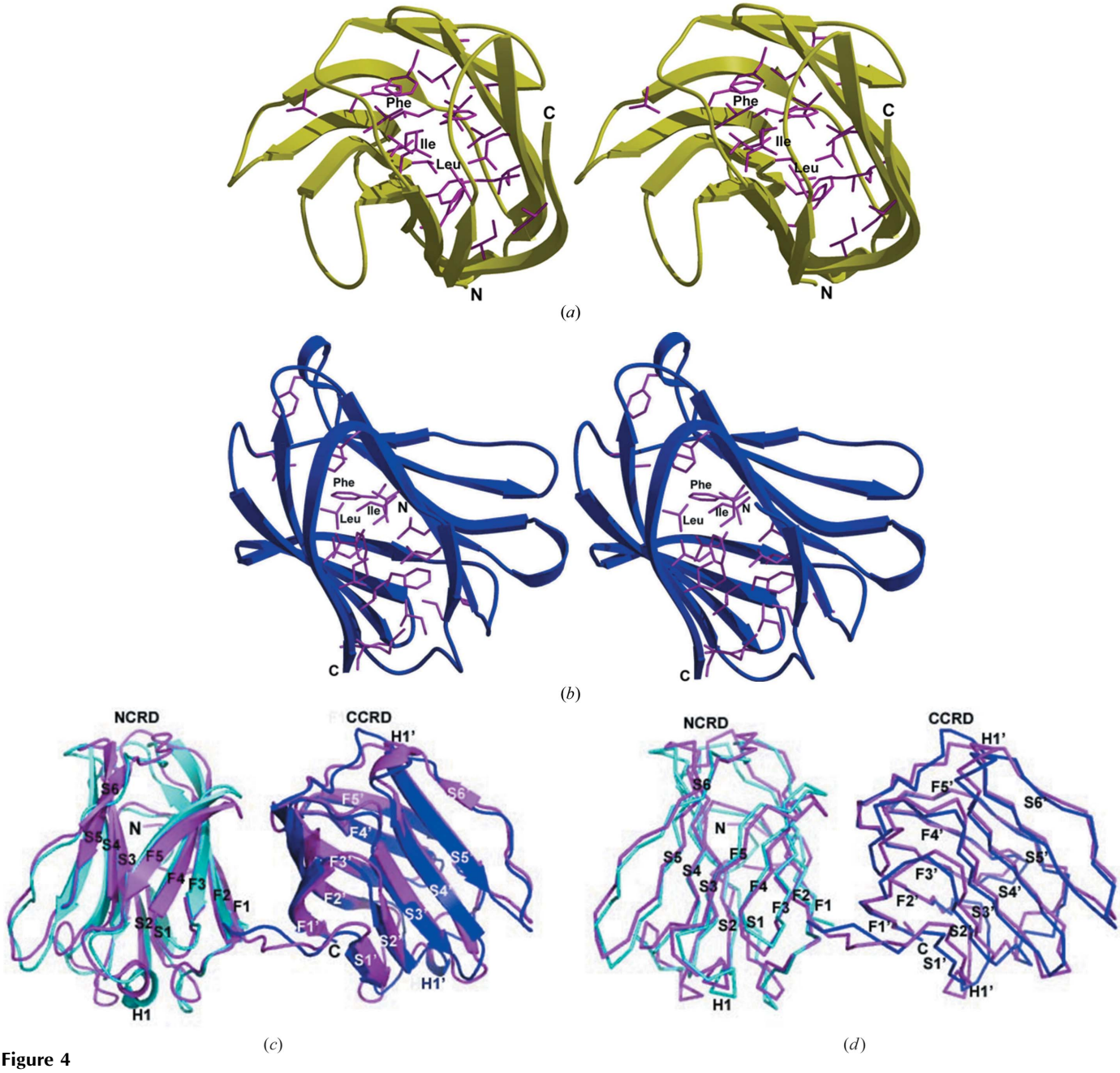


Figure 4

The hydrophobic environment and the superposition of the NCRD and CCRD in Tl-gal. (*a*) The NCRD hole surrounded by β-strands; hydrophobic residues are shown in purple. (*b*) The CCRD hole surrounded by β-strands; hydrophobic residues are shown in purple. (*c*) Superposition of a ribbon representation of the Tl-gal structure (residues 1–278; purple) on those of human galectin-7 (residues 1–135; cyan) and human galectin-9 (residues 188–323; blue). The regions with large conformational changes are found in loops. The β-strand regions with small conformation changes are shown and the α-helices are labelled H1 and H1′, respectively. (*d*) Superposition of the C^{α} trace of the Tl-gal structure (red) on those of human galectin-7 (cyan) and human galectin-9 (blue). A detailed view of the superposition of the β-strands is shown.

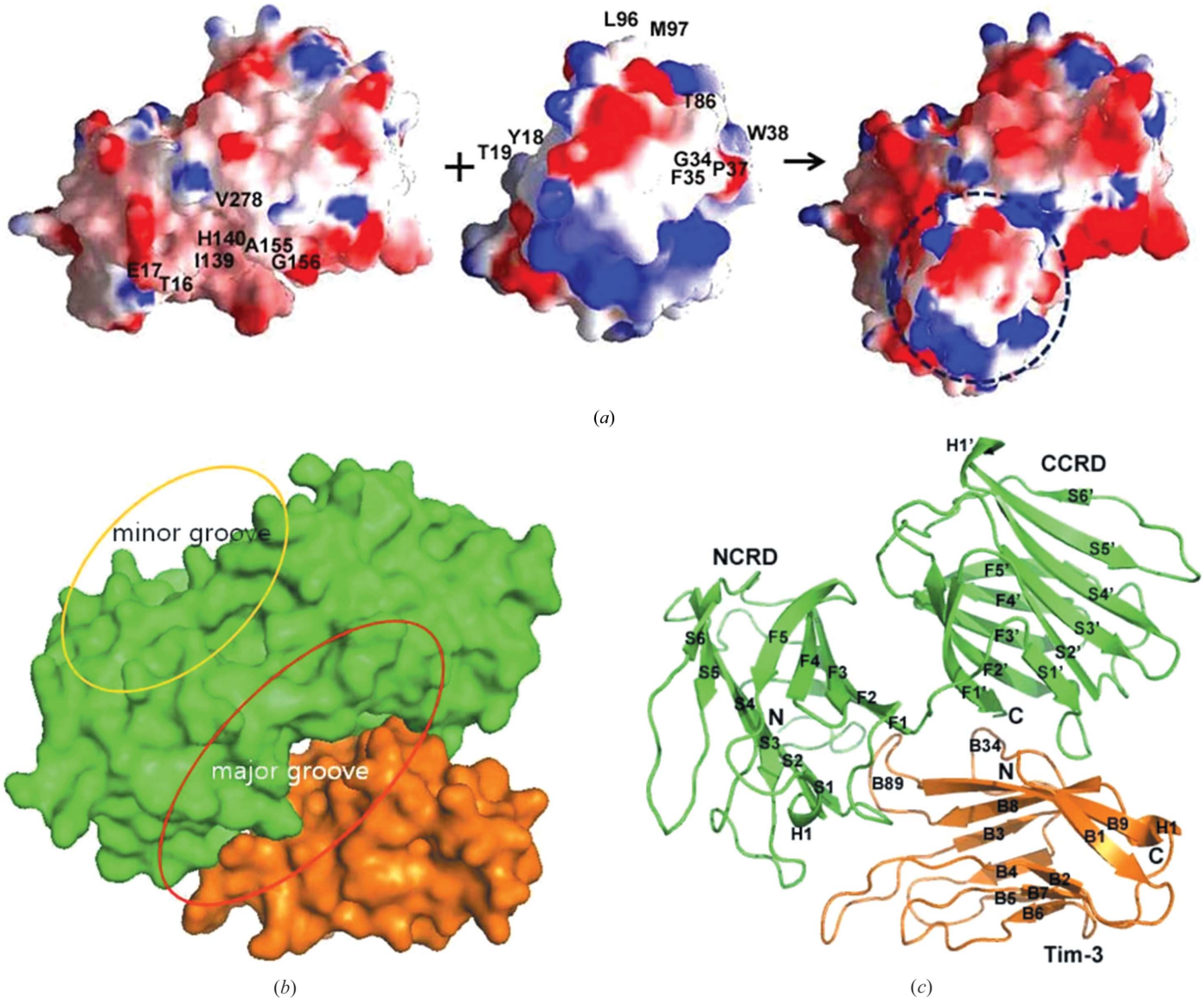


Figure 5
Model of the complex of Tl-gal and Tim-3. (*a*) Surface representations of Tl-gal (left panel), Tim-3 (middle panel) and the complex of Tl-gal with Tim-3 (right panel). The regions of Tl-gal that interact with Tim-3 are shown in the black dashed circle (right panel). The relative distribution of the surface charge is shown with acidic regions in red, basic regions in blue and neutral regions in white. (*b*) Surface representation of Tl-gal in green and Tim-3 in orange in the complex structure. The binding sites consist of a major groove outlined by the red circle and a minor groove outlined by the yellow circle. Tim-3 (shown in orange) interacts with the major groove. (*c*) Ribbon diagram of Tl-gal (green) and Tim-3 (orange).

In the present study, we have determined the first crystal structure of full-length *T. leonina* galectin (Tl-gal). The crystal structure shows that the NCRD and CCRD of Tl-gal are composed of six-stranded β-sheets and five-stranded β-sheets with a short α-helix, which together form a β-sandwich arrangement. The NCRD of Tl-gal resembles that of human galectin-7 and its CCRD resembles human galectin-9, but the residues in the interface and loop regions of the NCRD and the CCRD are flexible and related to interaction. To investigate the binding site of Tl-gal, we modelled the interaction between Tl-gal and Tim-3 immunoglobulin variable domain. Tim-3 is docked into a major groove of the Tl-gal structure, which is larger and deeper than the minor groove. The side chains and main chains of the loop region in the Tim-3 molecule are involved in the major groove. Tl-gal may function as a host galectin-9, thus functioning as a regulatory molecule in the host immune system. Structural information regarding Tl-gal will be useful in the development of novel therapeutic strategies to treat autoimmune diseases, inflammatory processes and allergic reactions.

We thank Drs Sung Chul Ha and Kyung-Jin Kim at the Pohang Light Source (PLS-7A), Republic of Korea for assistance with the X-ray diffraction experiments. This study was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012-0007607 to SBJ and 2012-0007941 to MSJ). This work was supported by

the 2011 Post-Doc. Development Program of Pusan National University.

References

Bidon-Wagner, N. & Le Pennec, J.-P. (2004). *Glycoconj. J.* **19**, 557–563.

Blois, S. M. *et al.* (2007). *Nature Med.* **13**, 1450–1457.

Boscher, C., Dennis, J. W. & Nabi, I. R. (2011). *Curr. Opin. Cell Biol.* **23**, 383–392.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Cao, E., Zang, X., Ramagopal, U. A., Mukhopadhaya, A., Fedorov, A., Fedorov, E., Zencheck, W. D., Lary, J. W., Cole, J. L., Deng, H., Xiao, H., Dilorenzo, T. P., Allison, J. P., Nathenson, S. G. & Almo, S. C. (2007). *Immunity*, **26**, 311–321.

DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.

Hirabayashi, J. & Kasai, K. (1993). *Glycobiology*, **3**, 297–304.

Hughes, R. C. (2004). *Glycoconj. J.* **19**, 621–629.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kellis, J. T., Nyberg, K., Sali, D. & Fersht, A. R. (1988). *Nature (London)*, **333**, 784–786.

Kim, J.-Y., Cho, M. K., Choi, S. H., Lee, K. H., Ahn, S. C., Kim, D.-H. & Yu, H. S. (2010). *Mol. Biochem. Parasitol.* **174**, 53–61.

Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Leonidas, D. D., Vatzaki, E. H., Vorum, H., Celis, J. E., Madsen, P. & Acharya, K. R. (1998). *Biochemistry*, **37**, 13930–13940.

Matsushita, N., Nishi, N., Seki, M., Matsumoto, R., Kuwabara, I., Liu, F.-T., Hata, Y., Nakamura, T. & Hirashima, M. (2000). *J. Biol. Chem.* **275**, 8355–8360.

Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Nichols, B. P. & Yanofsky, C. (1979). *Proc. Natl Acad. Sci. USA*, **76**, 5244–5248.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rabinovich, G. A., Baum, L. G., Tinari, N., Paganelli, R., Natoli, C., Liu, F. T. & Iacobelli, S. (2002). *Trends Immunol.* **23**, 313–320.

Shortle, D., Stites, W. E. & Meeker, A. K. (1990). *Biochemistry*, **29**, 8033–8041.

Su, A. I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K. A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., Cooke, M. P., Walker, J. R. & Hogenesch, J. B. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 6062–6067.

Than, N. G. *et al.* (2009). *Proc. Natl Acad. Sci. USA*, **106**, 9731–9736.

Than, N. G., Romero, R., Kim, C. J., McGowen, M. R., Papp, Z. & Wildman, D. E. (2012). *Trends Endocrinol. Metab.* **23**, 23–31.

Türeci, O., Schmitt, H., Fadle, N., Pfreundschuh, M. & Sahin, U. (1997). *J. Biol. Chem.* **272**, 6416–6422.

Yoshida, H., Teraoka, M., Nishi, N., Nakakita, S., Nakamura, T., Hirashima, M. & Kamitori, S. (2010). *J. Biol. Chem.* **285**, 36969–36976.

Zhu, C., Anderson, A. C., Schubart, A., Xiong, H., Imitola, J., Khoury, S. J., Zheng, X. X., Strom, T. B. & Kuchroo, V. K. (2005). *Nature Immunol.* **6**, 1245–1252.