



# Four crystal structures of human LLT1, a ligand of human NKR-P1, in varied glycosylation and oligomerization states

Tereza Skálová,<sup>a,\*</sup> Jan Bláha,<sup>b</sup> Karl Harlos,<sup>c</sup> Jarmila Dušková,<sup>a</sup> Tomáš Koval',<sup>d</sup> Jan Stránský,<sup>a</sup> Jindřich Hašek,<sup>a</sup> Ondřej Vaněk<sup>b</sup> and Jan Dohnálek<sup>a,d,\*</sup>

Received 18 September 2014

Accepted 22 December 2014

**Keywords:** LLT1; C-type lectin-like ligand.

**PDB references:** human LLT1, dimeric form, 4qkh; 4qki; hexameric form, 4qkj; monomeric form, 4qkg

**Supporting information:** this article has supporting information at journals.iucr.org/d

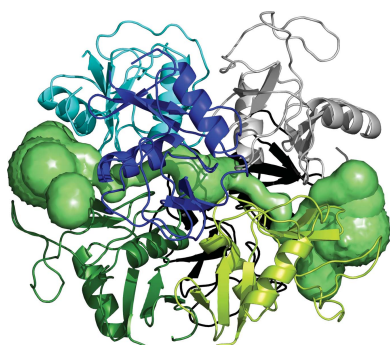
<sup>a</sup>Institute of Biotechnology, Academy of Sciences of the Czech Republic, v.v.i., Vídeňská 1083, 142 20 Praha 4, Czech Republic, <sup>b</sup>Department of Biochemistry, Faculty of Science, Charles University Prague, Hlavova 8, 128 40 Praha, Czech Republic, <sup>c</sup>Division of Structural Biology, The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, England, and <sup>d</sup>Institute of Macromolecular Chemistry, Academy of Sciences of the Czech Republic, v.v.i., Heyrovského nám. 2, 162 06 Praha 6, Czech Republic. \*Correspondence e-mail: t.skalova@gmail.com, dohnalek007@gmail.com

Human LLT1 is a C-type lectin-like ligand of NKR-P1 (CD161, gene *KLRB1*), a C-type lectin-like receptor of natural killer cells. Using X-ray diffraction, the first experimental structures of human LLT1 were determined. Four structures of LLT1 under various conditions were determined: monomeric, dimeric deglycosylated after the first *N*-acetylglucosamine unit in two forms and hexameric with homogeneous GlcNAc<sub>2</sub>Man<sub>5</sub> glycosylation. The dimeric form follows the classical dimerization mode of human CD69. The monomeric form keeps the same fold with the exception of the position of an outer part of the long loop region. The hexamer of glycosylated LLT1 consists of three classical dimers. The hexameric packing may indicate a possible mode of interaction of C-type lectin-like proteins in the glycosylated form.

## 1. Introduction

Natural killer (NK) cells are innate immune lymphocytes that possess the ability to recognize and induce the death of a broad range of target cells without prior antigen sensitization, including tumour, virally infected or stressed cells. Apart from direct cell-mediated cytotoxicity, NK cells also participate in the initiation and development of the adaptive immune response through the production and secretion of cytokines (Caligiuri, 2008; Vivier *et al.*, 2011). Their function is regulated by a fine balance of signals induced by the interaction of a vast array of surface-activating and surface-inhibitory receptors with ligands on the surface of target cells (Lanier, 2008; Vivier *et al.*, 2008; Bartel *et al.*, 2013). The activation of NK cells and the killing of the target cell is triggered when the expression of cellular health markers (MHC class I-like glycoproteins) that engage the inhibitory NK receptors is too low ('missing-self' recognition mode) or when the ligands for activating NK receptors are upregulated, usually in virally infected, stressed or malignant cells ('induced-self' mode).

The NK cell receptors are divided into two main structural classes: the immunoglobulin and C-type lectin-like (CTL) superfamilies. The CTL family of NK cells encoded within the natural gene complex (human chromosome 12) comprises proteins related to C-type lectins that have lost the ability to bind carbohydrates but have instead gained the ability to



OPEN ACCESS

recognize protein ligands (Yokoyama & Plougastel, 2003; Zelensky & Gready, 2005; Grobárová *et al.*, 2013). While certain CTL receptors are known to bind proteins with the MHC class I-like fold (*e.g.* NKG2D, CD94/NKG2A and murine Ly49 receptors; Bartel *et al.*, 2013), the NKR-P1 receptor subfamily does not share this specificity. Although the ligands for some rodent NKR-P1 receptors are still unknown, it has been shown that many of them recognize genetically and structurally highly related Clr/Ocil CTL receptors from the *CLEC2* gene subfamily (Aust *et al.*, 2009; Kveberg *et al.*, 2009). It is now accepted that this unique system of CTL receptor–CTL ligand interaction represents an alternative form of ‘missing-self’ recognition (Bartel *et al.*, 2013).

Lectin-like transcript 1 (LLT1, gene *CLEC2D*) was identified as a physiological ligand of NKR-P1, the sole described representative of the human NKR-P1 subfamily (CD161, gene *KLRB1*; Aldemir *et al.*, 2005; Rosen *et al.*, 2005). LLT1 is expressed primarily on activated lymphocytes (NK cells, T cells and B cells) and antigen-presenting cells (macrophages and dendritic cells; Germain *et al.*, 2011). Six alternatively spliced transcripts of the *CLEC2D* gene have been identified, with isoform 1 (coding for LLT1) being the only one able to interact with NKR-P1 (Germain *et al.*, 2010). The engagement of NKR-P1 on the NK cell with LLT1 on the target cell inhibits NK cell cytotoxicity and IFN $\gamma$  production (Aldemir *et al.*, 2005; Rosen *et al.*, 2005, 2008) and contributes to NK self-tolerance in a similar way to the orthologous rodent NKR-P1B–Clr-b receptor–ligand pair (Voigt *et al.*, 2007; Fine *et al.*, 2010; Williams *et al.*, 2012). It has been shown that human glioblastoma exploits this mechanism by the upregulation of the surface expression of LLT1 to escape the immunological response (Roth *et al.*, 2007). On the other hand, LLT1 is upregulated in response to both microbial and viral stimuli, and stimulation of NKR-P1-expressing T cells promotes their activation, proliferation and cytokine secretion (Huarte *et al.*, 2008; Germain *et al.*, 2011; Satkunanathan *et al.*, 2014). Thus, LLT1–NKR-P1 signalization may provide a link between pathogen pattern recognition and lymphocyte activation and represents a system that regulates both the innate and the adaptive immune responses.

Recently, we have described the first three-dimensional structure of mouse Clr-g, a representative of the murine *CLEC2* family (Skálová *et al.*, 2012). Similarly to mouse Clr-g, human LLT1 is a type II transmembrane glycoprotein which consists of an N-terminal cytoplasmic chain, transmembrane and stalk regions and a C-terminal CTL domain with two predicted N-linked glycosylation sites. In this study, we present several structures of LLT1 under various conditions: monomeric (LLT1\_mono), dimeric deglycosylated after the first *N*-acetylglucosamine unit (LLT1\_D1 and LLT1\_D2) and dimeric (packed into hexamers) with homogeneous GlcNAc<sub>2</sub>Man<sub>5</sub> glycosylation (LLT1\_glyco). All structures originate from protein expressed in HEK293S GnTI<sup>−</sup> cells (Reeves *et al.*, 2002). The dimeric form follows the classical dimerization mode of human CD69 (Natarajan *et al.*, 2000; Llera *et al.*, 2001; Vaněk *et al.*, 2008; Kolenko *et al.*, 2009).

## 2. Materials and methods

### 2.1. Protein expression and purification

The lectin-like domain of LLT1 was produced in HEK293S GnTI<sup>−</sup> cells as described by Bláha *et al.* (2015). Briefly, the expression construct corresponding to the CTL extracellular domain of LLT1 (Gln72–Val191) with a His176 to Cys176 mutation was cloned into the pTT28 plasmid, a derivative of the pTT5 plasmid (Durocher *et al.*, 2002) with an altered cloning site (Bláha *et al.*, 2015). A suspension culture of HEK293S GnTI<sup>−</sup> cells (Reeves *et al.*, 2002) was transiently transfected with a 1:3 mixture of the expression plasmid and 25 kDa linear polyethyleneimine. The resulting protein contained an N-terminal secretion leader and a C-terminal His<sub>8</sub> tag. 5–7 d post-transfection, the secreted recombinant protein was purified from the harvested production medium by two-step chromatography using batch IMAC on Talon resin (Clontech) followed by size-exclusion chromatography on a Superdex 200 10/300 GL column (GE Healthcare) in 10 mM HEPES pH 7.5 with 150 mM NaCl and 10 mM NaN<sub>3</sub>. The final product was concentrated to 20 mg ml<sup>−1</sup> using an Amicon Ultra concentrator (10 kDa molecular-weight cutoff; Merck Millipore).

For deglycosylation, GST-fused Endo F1 (Grueninger-Leitch *et al.*, 1996) was added to the concentrated protein solution in a 1:200 weight ratio and incubated overnight at room temperature. The slight precipitate that developed after deglycosylation was dissolved by the addition of L-arginine–HCl to a concentration of 0.4 M before setting up the crystallization drops.

The C-type lectin-like ectodomain of LLT1 contains two of the three canonical disulfide bridges (Cys75–Cys86 and Cys103–Cys184) found in homologous CTL receptors. Multiple alignment analysis showed that the formation of the third canonical disulfide bridge is impaired by the absence of the sixth evolutionarily conserved Cys residue, which is substituted by His176 in the wild-type LLT1 sequence. Based on previous results (Kamishikiryo *et al.*, 2011), we decided to clone and express the extracellular part (Gln72–Val191) of LLT1 with the His176 to Cys176 mutation. This mutation dramatically improved the stability, homogeneity and yield of the product (Bláha *et al.*, 2015). Expression in HEK293S GnTI<sup>−</sup> cells and subsequent purification provided highly pure protein with artificially homogeneous (GlcNAc<sub>2</sub>Man<sub>5</sub>) N-linked glycosylation with a typical yield of 3 mg of purified protein per litre of production culture. The disulfide-bond pattern as well as the occupancy of both N-glycosylation sites was verified by mass spectrometry (Bláha *et al.*, 2015).

### 2.2. Crystallization and data collection

**2.2.1. LLT1\_mono.** The protein (diluted to 14 mg ml<sup>−1</sup> in 0.4 M arginine, 4 mM HEPES, 120 mM NaCl, 4 mM NaN<sub>3</sub> pH 7.5) was crystallized using the sitting-drop vapour-diffusion method. Drops (100 nl reservoir solution and 100 nl protein solution) were set up using a Cartesian Honeybee 961 robot (Genomic Solutions) at 294 K. The reservoir consisted of 2 M ammonium sulfate, 0.1 M sodium citrate pH 3.5. A crystal with

the shape of a hexagonal plate (dimensions of  $100 \times 100 \times 10 \mu\text{m}$ ) was cryoprotected by soaking in the reservoir solution with the addition of 25% glycerol. The diffraction data were measured on a split segment of the multi-crystal on beamline I02 of Diamond Light Source (DLS) using an ADSC Q315 CCD detector at 100 K.

**2.2.2. LLT1\_D1 and LLT1\_D2.** The protein was crystallized as above. The reservoir consisted of 30% (w/v) polyethylene glycol (PEG) 6000, 0.1 M HEPES pH 7.0. A cuboid-shaped crystal of dimensions  $\sim 60 \times 60 \times 120 \mu\text{m}$  was cryoprotected as above. The data were measured on beamline I04-1 at DLS using a PILATUS 2M detector at 100 K. Both data sets were collected using crystals from the same crystallization condition. Interestingly, both crystals belonged to the same space group,  $P2_12_12_1$ , but the crystals differed in solvent content (36% for LLT1\_D1 and 25% for LLT1\_D2) and in unit-cell parameters (a difference of 4 Å in  $b$  and 8 Å in  $c$ ).

**2.2.3. LLT1\_glyco.** Protein diluted to  $14 \text{ mg ml}^{-1}$  in 10 mM HEPES, 150 mM NaCl, 10 mM  $\text{NaN}_3$  pH 7.5 was crystallized using the hanging-drop vapour-diffusion method (with drops consisting of 1  $\mu\text{l}$  reservoir solution and 1  $\mu\text{l}$  protein solution) with reservoir consisting of 40% (v/v) PEG 300, 0.1 M citrate-phosphate buffer pH 4.2 at a temperature of 288 K. A rod-shaped crystal of dimensions  $\sim 200 \times 50 \times 50 \mu\text{m}$  was vitrified without cryoprotection. Diffraction data were measured on BM 14.1 of the BESSY II synchrotron-radiation source (Mueller *et al.*, 2012) at the Hemholtz-Zentrum Berlin using a MAR Mosaic 225 CCD detector at 100 K.

### 2.3. Structure determination

The data sets for LLT1\_mono and LLT1\_glyco were processed using the *HKL-2000* package (*DENZO* and *SCALEPACK*; Otwinowski & Minor, 1997). Intensities were converted to amplitudes and imported into the CCP4 format with *TRUNCATE* (French & Wilson, 1978). The data for LLT1\_D1 and LLT1\_D2 were processed in *MOSFLM* (Leslie & Powell, 2007) using the *iMosflm* interface (Battye *et al.*, 2011) and were scaled in *AIMLESS* (Evans & Murshudov, 2013). The data parameters are shown in Table 1.

The data for LLT1\_glyco showed strong anisotropy; therefore, the data processed in *DENZO* and *SCALEPACK* were then rescaled using the *Diffraction Anisotropy Server* (<http://services.mbi.ucla.edu/anisotscale/>; Strong *et al.*, 2006). The processed data comprised 4186 reflections. An isotropic  $B$  factor of  $-23.39 \text{ \AA}^2$  was applied to restore the magnitude of the high-resolution reflections diminished by anisotropic scaling.

The phase problem for the LLT1\_D1, LLT1\_D2 and LLT1\_glyco structures was solved by molecular replacement in *BALBES* (Long *et al.*, 2008; Keegan *et al.*, 2013) using the structure of human CD69 as a search model (Kolenko *et al.*, 2009; PDB entry 3hup). The structure of LLT1\_mono was solved in *MOLREP* (Vagin & Teplyakov, 2010) using one chain of LLT1\_glyco as a search model.

All of the structures were refined in *REFMAC5* (Murshudov *et al.*, 2011) with manual changes performed in

*Coot* (Emsley *et al.*, 2010). Structure parameters are shown in Table 1. After several cycles of rigid-body refinement, the structures were refined using restrained refinement including H atoms (not deposited in the Protein Data Bank). Non-crystallographic symmetry between two monomers of the dimer was applied in the cases of LLT1\_D1, LLT1\_D2 and LLT1\_glyco in the first steps of the refinement. The refinement was performed with the omission of 5% of the reflections (test set with  $R_{\text{free}}$  flag). In the case of LLT1\_glyco with a lower resolution and a lower number of reflections, 6% of the reflections (223 reflections) were used for the test set. The last cycle of refinement of all four structures was performed with all reflections including the test set.

The choice of the correct space group for LLT1\_glyco was ambiguous and structure refinement was tested in 16 space groups alternative to  $P2_1$  using *Zanuda* (Lebedev & Isupov, 2014) and then also manually in  $P2_1$ ,  $C222$ ,  $C222_1$ ,  $P6_3$  and  $P6_322$ . Finally,  $P6_322$  was chosen as the correct space group because lower symmetry space groups did not show apparent differences among the molecules in the asymmetric unit and refinement was less stable.

### 2.4. Structure quality: geometry and electron density

**2.4.1. LLT1\_mono.** According to the *MolProbity* Ramachandran plot (Chen *et al.*, 2010), 95% of the residues of the structure lie in the favoured regions. There is one outlier, Gln83. This is a residue in a loop with an alternative conformation of the main chain. The residue is in good agreement with the  $2mF_o - DF_c$  and  $mF_o - DF_c$  electron-density Fourier maps.

One sulfate anion is localized in the structure, bound to Lys181 N $^\epsilon$ , Ser129 N and Ser129 O $^\gamma$ . The peptide bond Lys126-Gly127 in the vicinity of the sulfate anion has *cis-trans* alternative conformations.

**2.4.2. LLT1\_D1.** This structure with high resolution has residues generally well localized in  $2mF_o - DF_c$  and  $mF_o - DF_c$  maps; however, the Fourier maps have lower quality in the region of the long loop (residues 137–160), mainly in chain A. According to the *MolProbity* Ramachandran plot, 98% of the residues lie in the favoured regions and there is only one outlier, residue AGln139, which is in a turn of a loop.

There are unmodelled peaks in the  $2mF_o - DF_c$  Fourier maps. The electron density for the side chain of Lys186 in both chains is elongated, which may indicate covalent modification of this lysine. Moreover, arginine as added to the LLT1 sample after deglycosylation to improve protein folding is probably localized on the protein surface, bound by hydrogen bonds to residues BAsp104 and BAsp109, but the electron density is not sufficiently clear to build the whole arginine residue.

Surprisingly, the conformation of the Gln83-Arg84 peptide bond differs significantly between chain A (Gln83 torsion angles  $\varphi = 47^\circ$ ,  $\psi = 51^\circ$ ) and chain B ( $\varphi = 54^\circ$ ,  $\psi = -122^\circ$ ). Electron density is well defined in both cases. In chain B the value lies in the allowed region of the Ramachandran plot and BGLn83 O forms a water-mediated hydrogen bond to ACys75 N. In chain A, the torsion-angle values lie in the

**Table 1**  
Data-collection statistics and structure-refinement parameters.

Values in parentheses are for the highest resolution shell.

	LLT1_mono	LLT1_D1	LLT1_D2	LLT1_glyco
PDB code	4qkg	4qkh	4qki	4qkj
pH of crystallization condition	3.5	7.0	7.0	4.2
Glycosylation	GlcNAc	GlcNAc	GlcNAc	GlcNAc <sub>2</sub> Man <sub>5</sub>
Oligomer	Monomer	Dimer	Dimer	Hexamer
Data-collection statistics				
Resolution range (Å)	50.0–1.95 (1.98–1.95)	47.3–1.8 (1.84–1.80)	43.7–1.8 (1.84–1.80)	39.0–2.7 (2.76–2.70)
Space group	<i>P</i> <sub>3</sub> <sub>2</sub> <sub>1</sub>	<i>P</i> <sub>2</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub>	<i>P</i> <sub>2</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub>	<i>P</i> <sub>6</sub> <sub>3</sub> <sub>2</sub>
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 47.3, <i>c</i> = 106.1	<i>a</i> = 50.9, <i>b</i> = 57.8, <i>c</i> = 82.3	<i>a</i> = 51.3, <i>b</i> = 54.1, <i>c</i> = 74.2	<i>a</i> = <i>b</i> = 70.1, <i>c</i> = 101.7
Radiation source	I02, DLS	I04-1, DLS	I04-1, DLS	BM 14.1, BESSY II
Detector	ADSC Q315 CCD	PILATUS 2M	PILATUS 2M	MAR Mosaic 225 CCD
Data-processing software	<i>HKL</i> -2000 [ <i>DENZO</i> , <i>SCALEPACK</i> ]	<i>MOSFLM</i> , <i>AIMLESS</i>	<i>MOSFLM</i> , <i>AIMLESS</i>	<i>HKL</i> -2000 [ <i>DENZO</i> , <i>SCALEPACK</i> ], anisotropy server
Wavelength (Å)	0.9796	0.9200	0.9200	0.9184
No. of images	220	900	1800	80 [images 1–50, 330–360]
Crystal-to-detector distance (mm)	256.1	325.2	251.9	288.1
Exposure time per image (s)	1	0.2	0.2	2
Oscillation width (°)	0.5	0.2	0.2	0.5
No. of observations	66670 (3340)	175131 (5869)	113353 (6521)	478731 (1201)
No. of unique reflections	10555 (506)	21104 (1087)	18996 (1050)	4431 (273)
Data completeness (%)	99.7 (100)	91.8 (81.0)	96.7 (91.6)	99.5 (99.3)
Average multiplicity	6.3 (6.6)	8.3 (5.4)	6.0 (6.2)	4.3 (4.4)
Mosaicity (°)	0.3–0.7	1.8	1.3	0.8–1.4
$\langle I/\sigma(I) \rangle$	21.2 (2.9)	11.3 (3.0)	11.8 (3.7)	16.2 (2.0)
Solvent content (%)	44	36	25	46
Matthews coefficient (Å <sup>3</sup> Da <sup>-1</sup> )	2.18	1.93	1.64	2.30
<i>B</i> factor from Wilson plot (Å <sup>2</sup> )	34.0	20.7	17.7	66.1
<i>R</i> <sub>merge</sub> <sup>†</sup>	0.074 (0.599)	0.093 (0.462)	0.082 (0.404)	0.096 (0.757)
<i>R</i> <sub>p.i.m.</sub> <sup>‡</sup>	0.032 (0.251)	0.045 (0.309)	0.053 (0.257)	0.031 (0.171)
Structure parameters				
<i>R</i> <sub>work</sub> <sup>§</sup> (%)	19.4	17.5	17.8	22.1
<i>R</i> <sub>free</sub> <sup>§</sup> (%)	24.7	25.0	24.4	29.8
<i>R</i> <sub>all</sub> <sup>§</sup> (%)	20.1	18.0	18.3	23.1
Average <i>B</i> factor (Å <sup>2</sup> )	39.8	27.6	20.9	40.1
R.m.s.d. from ideal bond lengths (Å)	0.016	0.016	0.016	0.017
R.m.s.d. from ideal bond angles (°)	1.9	1.7	1.8	2.0
No. of monomers per asymmetric unit	1	2	2	1
Amino-acid residues located	A72–A146, A161–A193	A70–A192, B72–B192	A73–A194, B73–B192	A74–A188
Asn residues with located GlcNAc	A95	A95, A147, B95, B147	A95, A147, B95, B147	A95
No. of water molecules	42	215	167	15
Other localized moieties	1 × SO <sub>4</sub> <sup>2-</sup>	—	—	—
Ramachandran statistics: residues in favoured region (%)	95	98	98	87

<sup>†</sup>  $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$  (Diederichs & Karplus, 1997), where  $I_i(hkl)$  and  $\langle I(hkl) \rangle$  are the observed individual and mean intensities of a reflection with indices  $hkl$ , respectively,  $\sum_i$  is the sum over the individual measurements of a reflection with indices  $hkl$  and  $\sum_{hkl}$  is the sum over all reflections. <sup>‡</sup>  $R_{\text{p.i.m.}} = \sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$  (Weiss, 2001), where  $N(hkl)$  is the redundancy of the reflection with indices  $hkl$ . <sup>§</sup>  $R_{\text{work}} = \sum_{hkl} | |F_{\text{obs}}| - |F_{\text{calc}}| | / \sum_{hkl} |F_{\text{obs}}|$ , where  $F_{\text{obs}}$  and  $F_{\text{calc}}$  are the observed and calculated structure-factor amplitudes for the reflection with indices  $hkl$  for the working set of reflections.  $R_{\text{free}}$  is the same as  $R_{\text{work}}$  but is calculated for 5–6% of the data omitted from refinement.  $R_{\text{all}}$  sums over all reflections.

favoured region of the Ramachandran plot but the connection formed between chains *A* and *B* is weaker: it is mediated by hydrogen bonds formed by two water molecules connecting AArg84 N and BCys75 N.

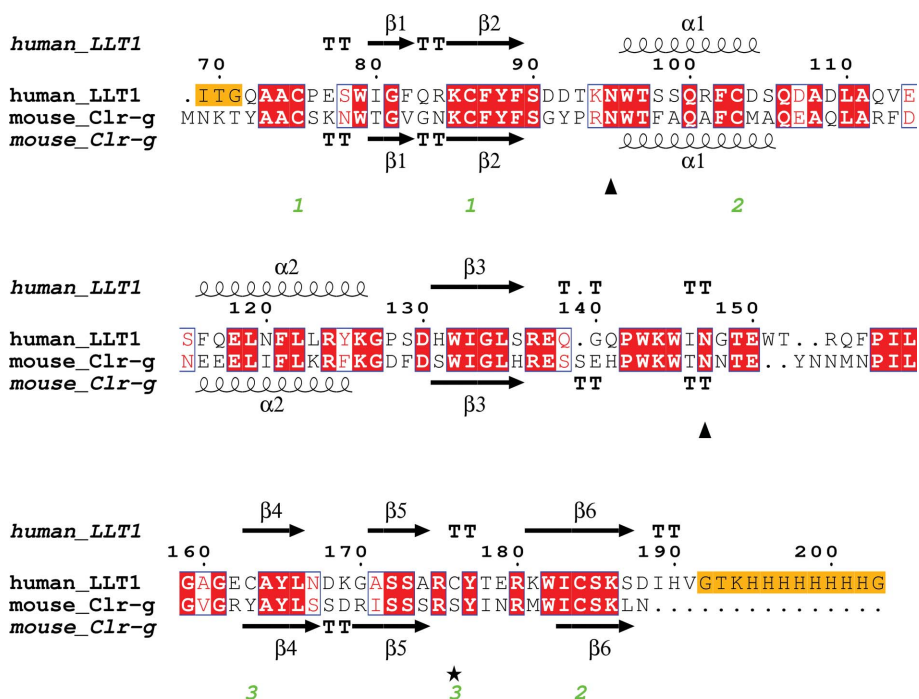
**2.4.3. LLT1\_D2.** Of the four presented structures, this structure has the best quality electron density because of its high resolution (1.8 Å) and low solvent content (25%). The long loop region of both chains is in contact with symmetry-related molecules in the crystal and is better localized in electron density than in the other presented structures. According to the *MolProbity* Ramachandran plot, 98% of residues lie in the favoured regions and there are no outliers.

Unlike in LLT1\_D1, there are no signs of Lys186 modification. Solvent arginine bound to BAsp104 and BAsp109

could be present, but is less apparent than in LLT1\_D1. The Gln83-Arg84 peptide bond is modelled in the same orientation in both chains, even though there are some signs of the alternative conformation in chain *A*.

**2.4.4. LLT1\_glyco.** The structure of glycosylated LLT1 has lower resolution (2.7 Å) and the data have an anisotropic character. According to the *MolProbity* Ramachandran plot, 87% of residues lie in the favoured regions and there are two outliers, ASer105 and AIle146, both of which are in turns of different loops.

Tyr88 has elongated electron density which may indicate covalent modification of this tyrosine. Some positive signal in the  $mF_o - DF_c$  map behind the hydroxyl group of Tyr88 is also present in chain *B* of LLT1\_D1.



**Figure 1**  
The LLT1 expression construct used for crystallization and its sequence alignment (*Clustal Omega*; Thompson *et al.*, 1994; Sievers *et al.*, 2011) with mouse Clr-g. The H176C mutation is denoted by an asterisk, glycosylated Asn residues are denoted by triangles and the three pairs of disulfide bonds in LLT1 are numbered. Uncleaved residues of the N-terminal signal sequence and the C-terminal tag are in orange boxes. Secondary structure was assigned by *DSSP* (Kabsch & Sander, 1983).

### 2.5. Electrostatic potential

The electrostatic potential of LLT1\_D2 was computed by solving the Poisson–Boltzmann equation in *APBS* (Baker *et al.*, 2001). H atoms were added in *PROPKA* (Li *et al.*, 2005), including optimization of hydrogen bonds for the protein at pH 7.5. Partial charges were assigned in *PDB2PQR* (Dolinsky *et al.*, 2007) based on the *AMBER* potential. The Poisson–Boltzmann equation was solved using values of the dielectric constants of  $\epsilon(\text{solvent}) = 78.54$  and  $\epsilon(\text{protein}) = 2$  and assuming a 0.225 M concentration of ions in solution with charges  $+e$  and  $-e$  and radius 2 Å.

### 2.6. Sedimentation analysis

The oligomeric state of the produced protein was analyzed in a ProteomeLab XL-I analytical ultracentrifuge equipped with an An-50 Ti rotor (Beckman Coulter, USA) using a sedimentation-velocity experiment. Samples of glycosylated LLT1 diluted with the buffer used in size-exclusion chromatography to the desired concentration were spun at 50 000 rev min<sup>-1</sup> at 20°C and 100 scans with 0.003 cm spatial resolution were recorded using absorbance optics at 280 nm. Buffer density and protein partial specific volume were estimated in *SEDNTERP* (<http://sednterp.unh.edu/>). The data were analyzed using *SEDPHAT* (Schuck, 2003) using the continuous size-distribution model and the monomer–dimer self-association model.

### 2.7. Dynamic light scattering

Measurement of dynamic light scattering was performed with a Zetasizer Nano (Malvern Instruments) and a 45 µl quartz cuvette with ~40 µl LLT1 solution (sample corresponding to LLT1\_glyco) diluted to 1 mg ml<sup>-1</sup> in a buffer solution consisting of 10 mM HEPES, 150 mM NaCl, 10 mM NaN<sub>3</sub> pH 7.5.

### 2.8. Figure preparation

Fig. 1 was prepared using *ESPrript* (<http://esprript.ibcp.fr>; Gouet *et al.*, 2003). Figs. 2–6 and Supplementary Fig. S1 were prepared in *PyMOL* (Schrödinger; <http://www.pymol.org>).

## 3. Results and discussion

The extracellular part of LLT1 was expressed in the HEK293S GnTI<sup>-</sup> cell line and was purified. Three crystal forms were grown at pH 3.5, 4.2 and 7.0 and led to four structures of LLT1: monomeric, dimeric (with altered crystal packing) and dimeric packed as hexamers. The hexameric structure of LLT1 has GlcNAc<sub>2</sub>Man<sub>5</sub> N-glycosylation; in the other three LLT1 structures the protein is deglycosylated beyond the first GlcNAc.

### 3.1. Overview of the crystal structures of LLT1

The crystallized extracellular part of human LLT1 corresponds to the sequence deposited in the UniProt database under code Q9UHP7, isoform 1, starting with Gln72 and with a His176 to Cys176 mutation (for the formation of a Cys163–Cys176 disulfide bond for greater stabilization of the protein fold) and with the addition of ITG (the remnant of a secretion leader) at the beginning of the chain and of a GTKHHHHHHHHHG tag at the end of the chain (Fig. 1). The calculated molecular weight of this construct is 15.7 kDa (18.1 kDa taking N-glycosylation into account).

The extracellular part of human LLT1 has the C-type lectin-like fold (Fig. 2*a*) with two  $\alpha$ -helices, two antiparallel  $\beta$ -sheets and three disulfide bonds (Cys70–Cys86, Cys103–Cys184 and Cys163–Cys176). The third (artificial) disulfide bond induced by the mutation of His176 to Cys176 is located in the same position as the third disulfide bond occurring naturally in CD69. The overall manner of dimerization of both its glycosylated and deglycosylated dimeric forms (produced in a human cell line in a close-to-native form) corresponds to the classical dimerization mode of human CD69 or of mouse Clr-g (one of the mouse orthologues of LLT1) both produced in bacteria and refolded from inclusion bodies. The N- and C-termini of both chains in this type of dimer are localized on

one side (Fig. 2*b*) and their position allows protein anchoring in the cell membrane, which supports the biological relevancy of this dimer.

The protein used to obtain the LLT1\_glyco structure contained homogeneous GlcNAc<sub>2</sub>Man<sub>5</sub> glycosylation, while the protein used to obtain the LLT1\_mono, LLT1\_D1 and LLT1\_D2 structures was deglycosylated beyond the first GlcNAc. However, only the first GlcNAc unit could be modelled in LLT1\_glyco. The subsequent oligosaccharide moieties cannot be modelled but are visible at low electron-density levels.

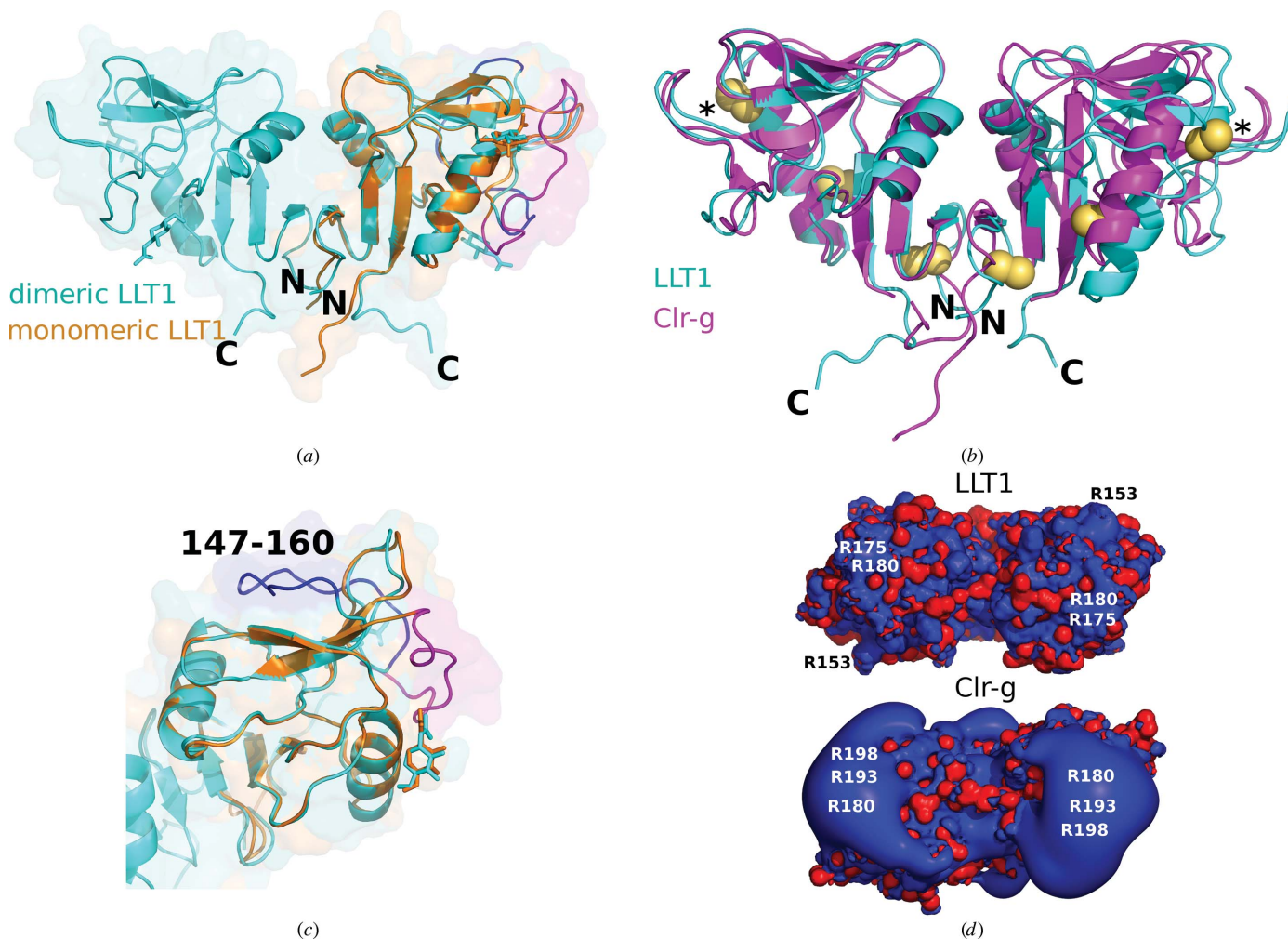
In LLT1\_mono, residues 147–160 (the outer loop of the long loop region) are poorly visible and were not built in the deposited structure. However, the approximate position of the loop is apparent in the electron density and differs significantly from its position in LLT1 in its dimeric form (Figs. 2*c* and 3*a*).

The structure of LLT1\_mono with the approximate position of the loop is available from the authors upon request.

### 3.2. Comparison of monomeric and dimeric LLT1

Superposition of chain *A* of LLT1\_D2 on LLT1\_mono by secondary-structure matching in *Coot* gives an r.m.s. deviation of C<sup>α</sup> atoms of 0.6 Å. The core of the monomer and the dimer has the same structure, but there are some structural differences in the outer parts of the CTL domain.

As mentioned above, there is an important difference in the placement of the outer loop (residues 147–160) of the long loop region (residues 137–160). While in LLT1\_D2 this loop lies in the direction of β-sheet 171–174 and α-helix 116–126 of the same chain, as is common in the structures of similar proteins (*e.g.* the CD69 dimer), in LLT1\_mono the loop is



**Figure 2**

(*a*) Structures of the dimeric (LLT1\_D2, cyan) and monomeric (LLT1\_mono, orange) forms of LLT1. The loop is in its standard dimeric position in LLT1\_D2 (blue) and is in the swapped position in LLT1\_mono (magenta). The N- and C-termini of LLT1\_D2 are denoted. (*b*) Comparison of the structure of LLT1\_D2 (cyan) with its mouse orthologue Clr-g (PDB entry 3rs1, magenta). The superposition was performed in *Coot* by SSM of chains *A* (on the left). The N- and C-termini of LLT1\_D2 are denoted. Disulfide bonds are shown as yellow spheres, with the artificial disulfide bonds induced by the H176C mutation denoted by asterisks. (*c*) A detailed view of the loop position discussed in (*a*). (*d*) Electrostatic equipotential surfaces of LLT1\_D2 (top) and mouse Clr-g (bottom). They are displayed at the 3 *kT/e* (blue) and -3 *kT/e* (red) levels as computed in *APBS* (Baker *et al.*, 2001). The view is from the 'top' of the dimer, *i.e.* from the termini-distal side.

turned to the opposite side with regard to the inner loop 135–145: it is situated near to  $\alpha$ -helix 96–104 of the same chain and makes contact with  $\alpha$ -helix 120–125 of a neighbouring monomer in the crystal (Figs. 2c and 3a).

There are also other structural differences between the monomer and dimer structures: the C-terminal part of the main chain, residues 190–194, is placed differently (Fig. 2a) and residues around Gln139 are mutually shifted; the distance between the C $^{\alpha}$  atoms in the monomeric and dimeric forms is 1.9 Å.

### 3.3. Dimer interface in dimeric LLT1

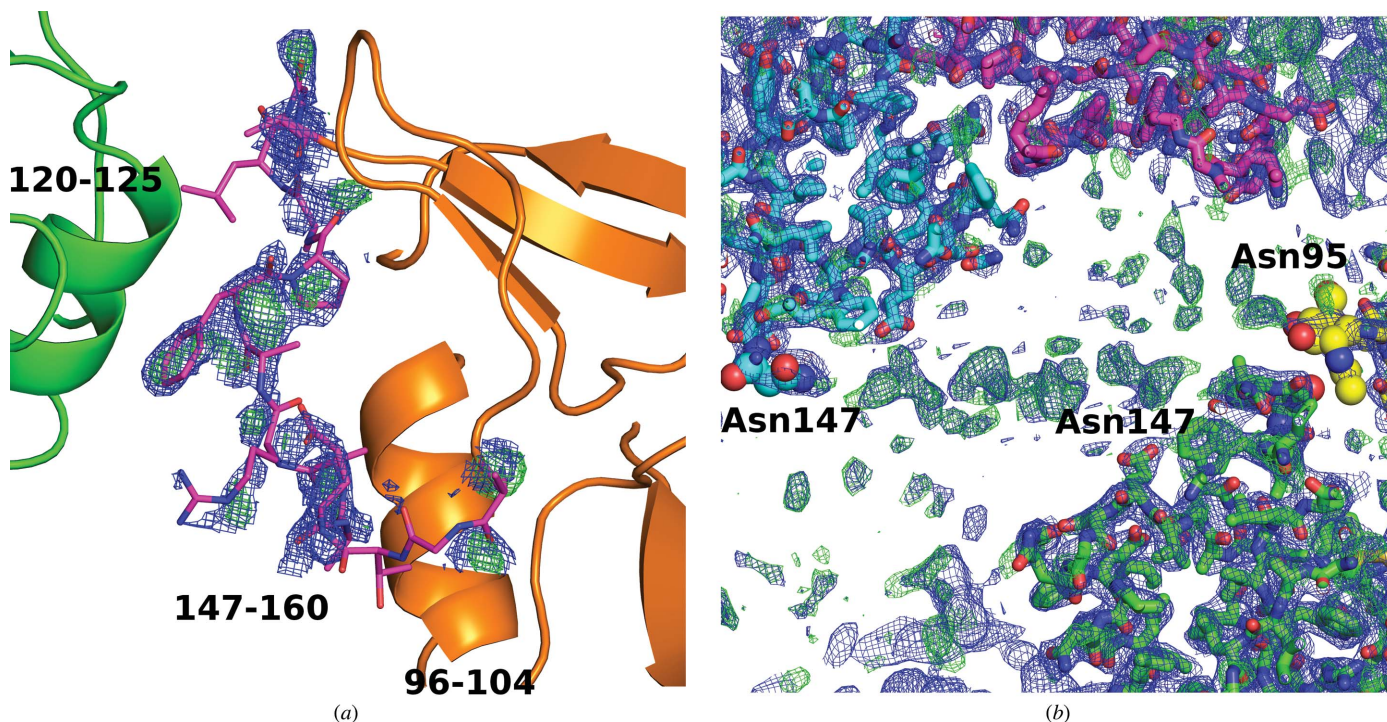
LLT1\_glyco, LLT1\_D1 and LLT1\_D2 share essentially the same structure at the level of the dimer; however, they differ slightly in the mutual orientation of the monomers within a dimer. This difference is especially apparent when the dimers are overlapped by only one chain: the rotation necessary to superimpose chain A onto chain B of the selected protein is exactly 180° for the LLT1\_glyco dimer crystallized in space group  $P6_322$ , in which the chains are related by a crystallographic rotation axis. In contrast, it is 173.7° for LLT1\_D1 and 177.7° for LLT1\_D2, both of which crystallized in space group  $P2_12_12_1$  with one dimer in the asymmetric unit. Owing to these differences in mutual orientation, there are differences in the area of the dimer interface (obtained using the PISA server; Krissinel & Henrick, 2007): the size of the contact area is

530 Å<sup>2</sup> for LLT1\_glyco (8% of the monomer surface area), 740 Å<sup>2</sup> for LLT1\_D2 (10% of the monomer surface area) and 810 Å<sup>2</sup> for LLT1\_D1 (11% of the monomer surface area).

The dimeric interface in all three LLT1 dimers is based on a hydrophobic core formed by phenyl rings, analogous to that observed in similar CTL dimers: Phe121, Tyr125, Phe89, Phe87 and Phe82 (Fig. 4). Additionally, there is an interaction formed by  $\pi$ -stacking of arginine residue Arg124 of both chains, and a pair of His190 residues forms a partial stacking interaction in the termini region (the part of the CTL domain oriented towards the cell membrane).

There are six common hydrogen bonds that connect the monomers in all three dimers: AArg124 N<sup>ε</sup>–BTyr125 O, AArg124 N<sup>η1</sup>–BLys126 O and the three analogous hydrogen bonds with exchanged chains. Additionally, there are some hydrogen bonds specific to individual cases: one additional hydrogen bond in the dimeric interface of LLT1\_glyco, two hydrogen bonds in LLT1\_D2 and six in LLT1\_D1.

This variability of the mutual chain orientation has been discussed previously for similar CTL protein dimers (CD94/NKG2A, Clr-g and CD69; Sullivan *et al.*, 2007; Skálová *et al.*, 2012). Considering the flexibility of the dimer interface, it seems that the flexible hydrophobic interactions are the most important contact for the existence of the dimer, its stability and its ability to bind a protein partner. These CTL dimers differ in their main purpose, which is surface partner binding,

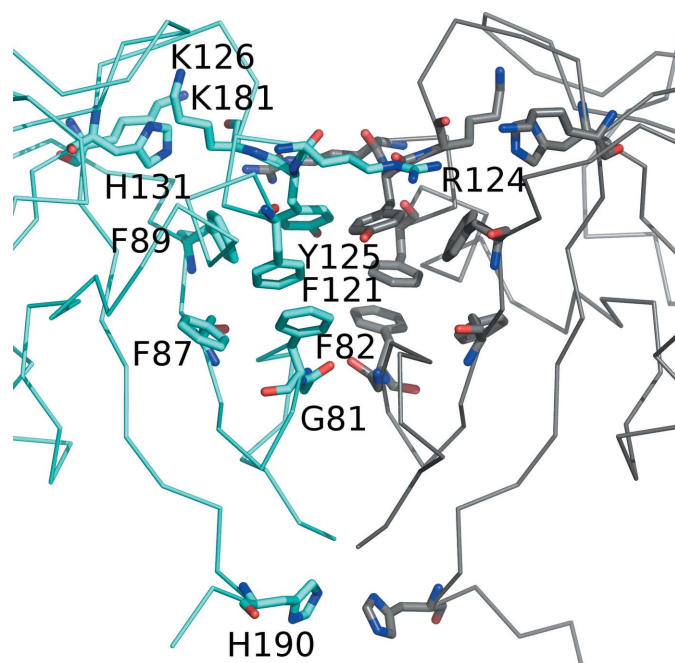


**Figure 3**  
 (a) Electron-density OMIT maps for the flexible loop (residues 147–160) placed in an unusual position in the monomeric structure of LLT1 (LLT1\_mono).  $2mF_o - DF_c$  (blue,  $0.8\sigma$ ) and  $mF_o - DF_c$  (green,  $2.5\sigma$ ) Fourier maps are drawn up to a distance of 1.6 Å from the atoms of residues 147–160. The residues are better localized in the middle of the loop, where they are stabilized by the nearby  $\alpha$ -helix of a neighbouring molecule in the crystal (green). The loop is not modelled in the deposited structure because weak signal in this region causes unstable refinement. (b) Two neighbouring hexamers in the crystal of LLT1\_glyco are connected by weak electron density belonging to the glycan chains. The glycosylated residues and GlcNAc units are shown as spheres.  $2mF_o - DF_c$  (blue,  $0.7\sigma$ ) and  $mF_o - DF_c$  (green,  $1.8\sigma$ ) Fourier maps are shown.

from for example the small aspartic proteases (such as HIV-1 protease), which in contrast require a strictly defined binding site at the dimer interface providing an environment for a specific peptide-bond cleavage. Increased flexibility within the dimer interface would be destructive for the function of proteases, unlike in this case of CTL protein–CTL ligand complexation, where the flexibility and ability to adapt the dimer shape to enable protein partner binding could be an advantage.

### 3.4. Crystal packing of monomeric LLT1

The crystal packing of LLT1\_mono was investigated to confirm that the form of the protein in this structure is really monomeric and that no possible partner participating in oligomer formation can be found among the symmetry-related copies. The crystal packing is shown in Fig. 5(a). The classical dimerization interface is left open to the solvent, without any crystal contacts. The largest intermolecular interface in the crystal (according to the *PISA* server; Krissinel & Henrick, 2007) has a surface area of 724 Å<sup>2</sup> (11% of the monomer surface area) and includes ten hydrogen bonds. The following interface is much weaker (372 Å<sup>2</sup>). The largest intermolecular contact is formed by the C- and N-termini of the chain with the region above the  $\beta$ -sheet of the neighbouring molecule where the outer loop of the long loop region is usually localized. In this structure, the outer loop (147–160) of the long loop region is turned over towards the other side along the inner loop and the terminal region of a neighbouring molecule takes its place (Figs. 2c and 5a). This contact with the C- and N-termini cannot be a biologically relevant contact, because in the cell the N-terminal chain continues to the membrane. This leads to the conclusion that in this structure the LLT1 monomer does not form any biologically relevant contacts.



**Figure 4**  
A detailed view of the LLT1 dimer interface (structure LLT1\_D2).

### 3.5. Crystal packing of dimeric LLT1

The strongest dimer–dimer contact in the LLT1\_D1 and LLT1\_D2 crystals is the contact of the turns of three loops (around residues 139, 160 and 177) with a ‘binding pocket’ at the dimer interface (in the part distant from the N- and C-termini; Fig. 5b). The contact has an area of 620 Å<sup>2</sup> and includes 15 hydrogen bonds. An *N*-acetylglucosamine unit bound to Asn95 is also in this region and is close to the ‘binding pocket’. This type of crystal packing would be much more complicated or impossible in the case where the protein was not deglycosylated after the first GlcNAc unit.

In our structural study of the Clr-g receptor (Skálová *et al.*, 2012), the strongest contact in the crystal was also connected with a bond to this ‘binding pocket’ at the dimer interface. In that case, it was the truncated N-terminus of the Clr-g chain making a tight interaction with the neighbouring dimer interface (the ‘binding pocket’), a clearly strong but biologically irrelevant interaction.

### 3.6. Role of pH in the formation of dimers

The monomeric form of LLT1 was crystallized at pH 3.5 and the changed protonation state at low pH corresponds with the changed preferred intermolecular contacts and the formation of oligomers.

As an example, residue His190, localized near the C-terminus in the dimerization interface, is charged at low pH and destabilizes this part of the interface. It is apparent from the structures that His190 is neutral in LLT1\_D1 and LLT1\_D2 (pH 7.0) and is close to His190 (3.1 and 3.3 Å) from the opposite chain, participating in a partial stacking interaction. In LLT1\_glyco (pH 4.2) His190 is not localized in electron density. In LLT1\_mono His190 is very likely to be protonated as it forms a hydrogen bond to the N-terminal Gln72 O of the same chain.

Residues His131 and Lys181 are found in LLT1\_D1 and LLT1\_D2 close (with regard to interaction of charges, ~5 Å) to the site in which the guanidinium group of Arg124 reaches over from the opposite chain to ‘lock’ the dimer interface, which is otherwise formed mainly by the hydrophobic core. Arg124 stacks on the side chain of Tyr125 and forms additional hydrogen bonds. In the low-pH structure LLT1\_mono, His131 is charged and this site attracts an SO<sub>4</sub><sup>2-</sup> ion. Under these conditions the positive charge of Lys181 and His131 repels Arg124 from its standard position and thus contributes to dimer destabilization.

As LLT1 is expected to function as a dimer, very low pH values would disable its dimerization and its function as such. A tumour environment often shows decreased pH, but not such drastically lowered values, and a decrease of pH in some cases also leads to higher activity of cytotoxic cells. Given that we observe standard dimers at pH 4.2 in the LLT1\_glyco structure and only the extreme pH of 3.5 in LLT1\_mono leads to dimer disruption, this direct observation of behaviour at extremely low pH is not relevant to normal biological limits.



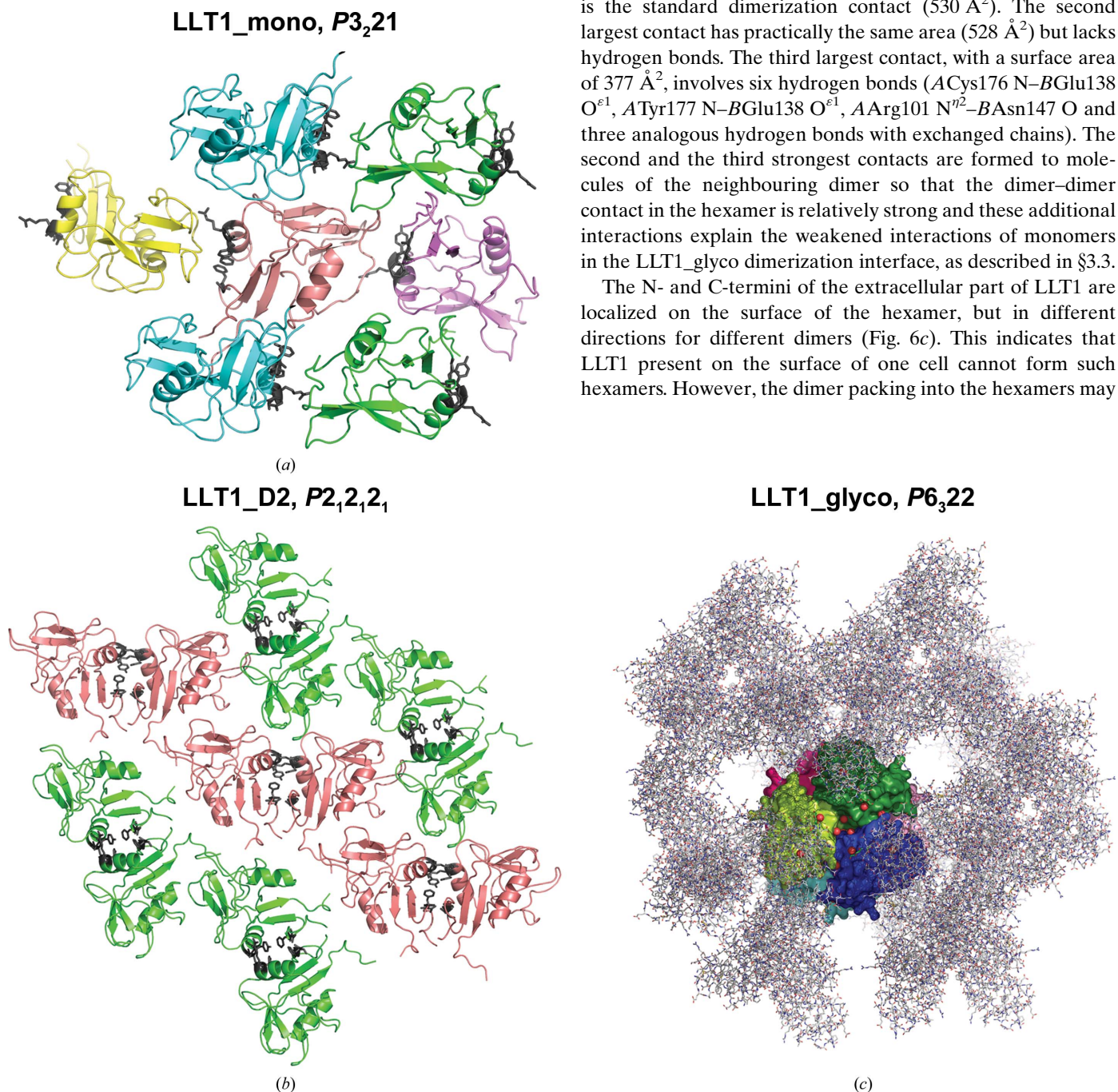
### 3.7. Hexamers of LLT1 with GlcNAc<sub>2</sub>Man<sub>5</sub> glycosylation

The LLT1<sub>glyco</sub> structure is packed into hexamers (Figs. 5c and 6). The hexamer is formed by three dimers, which are similar to those of LLT1<sub>D1</sub> and LLT1<sub>D2</sub>. The hexamer has one threefold and three perpendicular twofold axes of symmetry (point symmetry 32, where the asymmetric unit is one monomer). The hexamer of ~70 Å diameter contains a central cavity of about 10 Å in diameter. The cavity is formed

by the surface depressions of the three participating dimers found on the termini-distal surface. Six Pro128 residues are exposed into the cavity and six Asp168 and Lys169 residues with their side chains near to the proline residues together form the border of the cavity. The cavity is connected to the exterior of the hexamer by a channel along the threefold axis of the hexamer (Fig. 6b; the channel shape was computed in CAVER; Chovancova *et al.*, 2012).

The tightest monomer–monomer contact in the hexamer is the standard dimerization contact (530 Å<sup>2</sup>). The second largest contact has practically the same area (528 Å<sup>2</sup>) but lacks hydrogen bonds. The third largest contact, with a surface area of 377 Å<sup>2</sup>, involves six hydrogen bonds (ACys176 N–BGlu138 O<sup>ε1</sup>, ATyr177 N–BGlu138 O<sup>ε1</sup>, AArg101 N<sup>η2</sup>–BAsn147 O and three analogous hydrogen bonds with exchanged chains). The second and the third strongest contacts are formed to molecules of the neighbouring dimer so that the dimer–dimer contact in the hexamer is relatively strong and these additional interactions explain the weakened interactions of monomers in the LLT1<sub>glyco</sub> dimerization interface, as described in §3.3.

The N- and C-termini of the extracellular part of LLT1 are localized on the surface of the hexamer, but in different directions for different dimers (Fig. 6c). This indicates that LLT1 present on the surface of one cell cannot form such hexamers. However, the dimer packing into the hexamers may



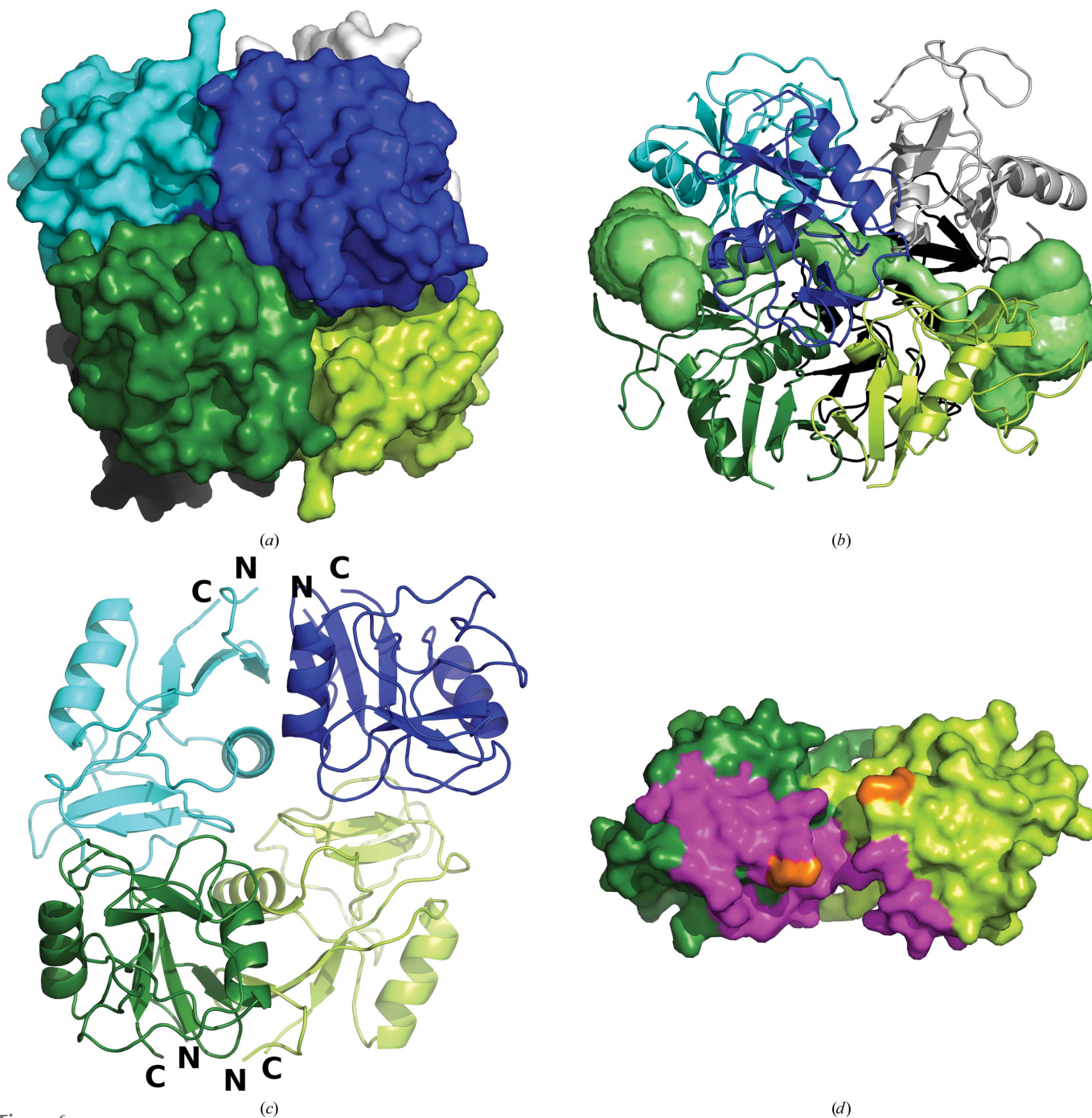
**Figure 5** Crystal packing of the presented LLT1 structures. (a) LLT1<sub>mono</sub>. Colours are according to the orientation of the protein molecule in the crystal. Residues that would form the dimer interface in the case of a dimer (Gly81, Phe82, Phe121, Arg124 and Tyr125) are shown as black sticks. (b) LLT1<sub>D2</sub> with the same colour coding. The dimer interface is shown as black sticks. LLT1<sub>D1</sub> has similar packing as LLT1<sub>D2</sub>. (c) LLT1<sub>glyco</sub>. One hexamer, containing three dimers of LLT1, is represented by a molecular surface colour-coded by individual protein chains and viewed along the crystallographic threefold axis. Surrounding symmetry-related chains are represented as sticks.

indicate a possible method of NK receptor–ligand interaction between two cells (see §3.11).

### 3.8. Role of glycosylation in LLT1\_glyco

In the hexamer, the glycosylated Asn residues Asn95 and Asn147 are localized on its surface. Three Asn95 residues

related by the threefold axis are at a mutual distance of 18 Å and are placed relatively close to the central channel. Three Asn147 residues located ~18 Å from the Asn95 residues in the direction away from the central channel are distant from each other. The other three Asn95 and three Asn147 residues are in the same formation on the opposite side of the hexamer.



**Figure 6**

The fully glycosylated form of LLT1 forms hexamers in crystals which are an assembly of three classical dimers. The dimers are distinguished by colour (dark and light green, dark and light blue and black and white). (a) A general view of the hexamer. (b) Visualization of the tunnel and the central cavity of the hexamer (computed in *CAVER*). A view perpendicular to the threefold axis, which is the direction of the tunnel. (c) Dimer–dimer contacts in the hexamer. One dimer of the hexamer is omitted to show the mutual orientation of two dimers. The N- and C-termini of the chains are denoted. (d) Contact residues of a dimer with its neighbour in the hexamer. Residues forming the LLT1 dimer–dimer contact up to a distance of 5 Å are shown in magenta. Residue Lys169 (important for interaction in the NKR-P1–LLT1 complex) is coloured orange. The contact residues are localized on the termini-distal side. Interacting residues were identified using *NCONTACT* from the *CCP4* package (Winn *et al.*, 2011).

However, Asn147 is very close to Asn95 from a symmetry-related hexamer in the unit cell: the distance between Asn147 O and Asn95 O<sup>δ1</sup> from a symmetry-related molecule is 7.7 Å (Fig. 3*b*). In spite of the fact that it is not possible to model any unit of glycosylation well at Asn147 because the electron density is not sufficiently unambiguous, there is apparent continuous electron density between Asn147 and Asn147 from the neighbouring hexamer (the distance between the Asn147 N<sup>δ2</sup> atoms is 33 Å; Fig. 3*b*) with peaks of up to 1.3σ in 2*mF*<sub>o</sub> – *DF*<sub>c</sub> and 3.6σ in *mF*<sub>o</sub> – *DF*<sub>c</sub> maps, indicating that the oligosaccharide chains interconnect molecules and indeed the individual hexamers in the crystal. Thus, we conclude that we do not observe any role of glycans in the formation of the hexamer; however, it is evident that the glycans contribute to the packing of the hexamers into the crystal lattice.

### 3.9. Influence of the His176Cys mutation on the LLT1 structure

The mutated residue 176 is localized on the surface of LLT1 near the long loop region. It is distant from the N- and C-terminal parts and also does not lie in the dimerization interface. The additional S–S bond involving Cys176 stabilizes the protein fold. It is probable that the His176 form of LLT1 would have a longer flexible part of the loop than the flexible part 147–160 observed in the structure of the Cys176 variant. The residue lies in a crystal contact region in the dimeric and hexameric structures, and therefore the wild-type form of LLT1 with His176 would have different crystal packing.

### 3.10. Comparison of LLT1, CD69 and Clr-g structures

LLT1 shares relatively high three-dimensional structure similarity with other dimeric CTL proteins. All comparisons were performed with the LLT1\_D2 coordinates. According to SSM structure superposition with *PDBeFold* (<http://www.ebi.ac.uk/msd-srv/ssm/>), LLT1 is structurally most similar to human CD69 (PDB entry 1e8i; Llera *et al.*, 2001; 36% sequence identity). Structure similarity to mouse Clr-g (PDB entry 3rs1; Skálová *et al.*, 2012; 44% sequence identity) has similar values when comparing monomers and is a little lower when comparing dimers. The parameters of the structure superpositions are as follows: LLT1 monomer/CD69 monomer, *Q*-score<sup>1</sup> 0.82 and r.m.s. deviation 1.10 Å, aligned on 115 residues; LLT1 dimer/CD6 dimer, *Q*-score 0.67 and r.m.s. deviation 1.67 Å, aligned on 223 residues; LLT1 monomer/Clr-g monomer, *Q*-score 0.82 and r.m.s. deviation 0.96 Å, aligned on 116 residues; LLT1 dimer/Clr-g dimer, *Q*-score 0.55 and r.m.s. deviation 2.10 Å, aligned on 219 residues.

The comparison of the structure of LLT1\_D2 with that of mouse Clr-g is shown in Fig. 2(*b*). The superposition was performed in *Coot* by SSM of chains *A*. There is an apparent difference in the mutual orientation of the monomers in the

<sup>1</sup> *Q*-score is the default sorting criterion of structure similarity in *PDBeFold* and is defined as  $Q = N_{\text{align}} \times N_{\text{align}} / ([1 + (\text{r.m.s.d.}/R_0)^2] \times N_{\text{res1}} \times N_{\text{res2}})$ , where  $R_0 = 3$  Å,  $N_{\text{res1}}$  and  $N_{\text{res2}}$  are the numbers of residues in the overlapping structures and  $N_{\text{align}}$  is the number of overlapped residues.

dimer of LLT1 in comparison with mouse Clr-g. There are shifts in the distant part of the long loop region of up to 5 Å. Differences are also present in the N- and C-terminal parts owing to the distinct lengths of the LLT1 and Clr-g constructs. Differences between LLT1 and CD69 are found in the same regions and are of the same character as the differences between LLT1 and Clr-g (not shown).

Electrostatic equipotential surfaces of LLT1 and mouse Clr-g (Skálová *et al.*, 2012) are shown in Fig. 2(*d*). There are some positive patches in the part distant from the N- and C-termini in LLT1. A similar and much stronger effect is observed for mouse Clr-g. The large patch in Clr-g is formed mainly owing to the contributions of Arg180, Arg193 and Arg198, which correspond to Glu162, Arg175 and Arg180 in LLT1. The change from Arg180 to Glu162 may explain the weaker positive patches in the case of LLT1.

### 3.11. Hypotheses about NKR-P1–LLT1 complexation

The hexamer is formed by the packing of three ‘standard’ CTL dimers. To the best of our knowledge, no such molecular arrangement has been observed for CTL proteins. Each dimer in the hexamer interacts with two other dimers. The dimer–dimer interaction in the hexamer is analyzed in Figs. 6(*c*) and 6(*d*). Could this dimer–dimer interaction indicate the manner of NK CTL receptor–CTL ligand complex formation or other biologically relevant actions?

This interaction of LLT1–LLT1 dimers in the hexamer is not a ‘face-to-face’ interaction, as is classically expected for NK CTL receptor–CTL ligand interaction (Kamishikiryo *et al.*, 2011); rather, one monomer of the dimer binds to one dimer and the second monomer binds to another dimer of the hexamer. In this aspect it is similar to the model of the NKR-P1F–Clr-g complex (based on electrostatic complementarity) in our previous study (Skálová *et al.*, 2012).

The only experimentally determined structure of a CTL NK receptor–CTL ligand complex is the X-ray structure of the monomeric receptor NKp65–dimeric KACL ligand complex (PDB entry 4iop; Li *et al.*, 2013), where the binding mode is of the monomer–dimer type. The sequence identity of the CTL domain of LLT1 to the CTL domain of KACL is 49%, while that of LLT1 to NKp65 is 30%.

We have built several models of the NKR-P1–LLT1 complex. The modelling of NKR-P1 was performed in *SWISS-MODEL* (Biasini *et al.*, 2014) based on the crystal structure of mouse NKR-P1A (PDB entry 3m9z; Kolenko, Rozbeský, Vaněk, Kopecký *et al.*, 2011; 46% sequence identity). The crystal structure of mouse NKR-P1A (Kolenko, Rozbeský, Vaněk, Bezouška *et al.*, 2011; Kolenko, Rozbeský, Vaněk, Kopecký *et al.*, 2011) has extended flexible loops which form crystal contacts and most probably occupy a different position in solution (Sovová *et al.*, 2011). However, the identified receptor–ligand interaction pairs do not lie in the loops; therefore, it is possible to use the crystal structure as a modelling basis for our purposes.

The first model is based on the NKp65–KACL structure. Both the model of NKR-P1 and the structure of LLT1 were

superimposed on the structure of the NKp65–KACL complex to simulate a model of the NKR-P1–LLT1 complex (Supplementary Fig. S1a). In this rather crude model it is possible to query the interaction between Lys169 of LLT1 and Glu205 of NKR-P1, which were identified by Kamishikiryo and coworkers as the key interacting residues based on SPR experiments (Kamishikiryo *et al.*, 2011). It should be noted that these SPR experiments were performed on nonglycosylated LLT1 produced in *E. coli*. In this first model, these two residues are 10 Å away from each other.

In the second model we exchanged the positions of the receptor and the ligand, *i.e.* NKR-P1 was superimposed on KACL and the individual chains of LLT1 on the positions of the monomers of NKp65 (Supplementary Fig. S1b). In this model, residues Lys169 of LLT1 and Glu205 of NKR-P1 are in direct contact.

The third model is based on the LLT1–LLT1 dimer–dimer mutual position in the hexameric structure. The NKR-P1 model was superimposed on the closest LLT1 chain in the

hexamer with respect to a fixed LLT1 dimer (Supplementary Figs. S1c and S1d). In this model, the distance between LLT1 Lys169 and NKR-P1 Glu205 is 20 Å.

Based on the assumption of close positioning of Lys169 and Glu205, we conclude that the LLT1 arrangement in the hexameric packing is not a probable model for NK CTL receptor–CTL ligand interaction. Interestingly, the only model enabling direct contact of Lys169 and Glu205 residues is that which requires disruption of the LLT1 dimer into monomers. It remains to be shown experimentally in the future whether this is truly the case.

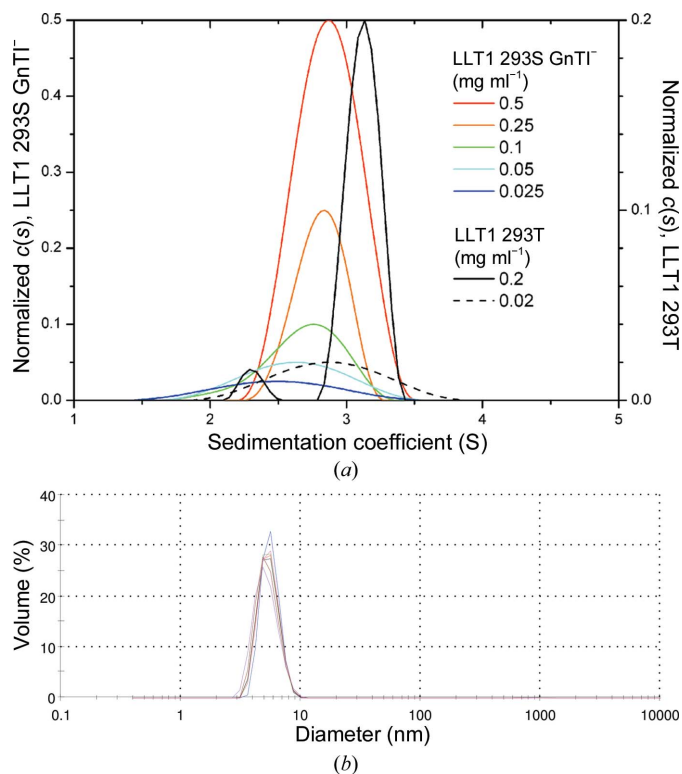
### 3.12. Oligomeric state of LLT1 in solution

Sedimentation-equilibrium analysis has previously shown that LLT1 forms noncovalent dimers in solution that do not depend on its N-glycosylation (Bláha *et al.*, 2015). However, to assess for the presence of either monomers or higher oligomers (*e.g.* hexamers), we performed sedimentation-velocity measurements. Firstly, LLT1 produced in the HEK293T cell line with wild-type complex N-glycosylation was analyzed. At 0.2 mg ml<sup>-1</sup> concentration, LLT1 behaves as two distinct species with  $s_{20,w}$  values of 2.43 and 3.14 ± 0.1 S corresponding to a monomer and a dimer, respectively (Fig. 7a, black line). Upon tenfold dilution the two peaks merged into one with an average  $s_{20,w}$  value of 2.87 S, pointing to a monomer–dimer equilibrium at low concentration (Fig. 7a, dashed line). No higher oligomers were detected. LLT1 with homogeneous GlcNAc<sub>2</sub>Man<sub>5</sub> N-glycosylation produced in the HEK293S GnTI<sup>-</sup> cell line behaved similarly, yielding broad size distributions that shifted towards lower sedimentation-coefficient values at lower protein concentrations (Fig. 7a, coloured lines), thus reflecting monomer–dimer equilibrium behaviour of the protein. The fact that monomeric and dimeric forms are not separated in these distributions might suggest that LLT1 produced with shorter N-glycans dimerizes more weakly and/or that equilibrium exchange is faster than for LLT1 with wild-type complex N-glycans, with an estimated  $K_d$  lying in the low micromolar range (~20 μM).

Dynamic light scattering was performed for LLT1 with homogeneous GlcNAc<sub>2</sub>Man<sub>5</sub> N-glycosylation produced in the HEK293S GnTI<sup>-</sup> cell line at 1 mg ml<sup>-1</sup> concentration (Fig. 7b). The experiment was repeated at temperatures of 291, 298 and 303 K. Particles with a diameter of 55 ± 10 Å were observed in all cases. This value corresponds to a Stokes diameter of 56 Å for the LLT1 dimer, as computed based on the three-dimensional structure of LLT1 in HYDROPRO (Ortega *et al.*, 2011), while the Stokes diameter computed for the LLT1 hexamer is 70 Å and that for the monomer is 20 Å.

## 4. Conclusions

This study has introduced the first X-ray structures of LLT1. A new conformation of the outer loop in the long loop region of a CTL receptor/ligand was observed; its newly observed position near α-helix 96–104 is probably the result of conformational sampling with satisfied packing positions of the



**Figure 7**  
(a) Sedimentation analysis of LLT1 oligomerization. Glycosylated LLT1 with homogeneous GlcNAc<sub>2</sub>Man<sub>5</sub> N-glycosylation produced in the HEK293S GnTI<sup>-</sup> cell line was characterized by a sedimentation-velocity experiment in an analytical ultracentrifuge at five different concentrations (coloured lines). The continuous size distribution of the sedimenting species is shown, which was normalized with respect to the difference in loading concentration. A broad size distribution that is shifted towards lower values with lower protein concentration points to monomer–dimer equilibrium; for comparison, data for LLT1 produced in the HEK293T cell line with wild-type complex N-glycosylation are shown (black and dashed lines), with separated signals for dimeric and monomeric protein at higher protein concentration. (b) Size distribution of LLT1 measured by dynamic light scattering and scaled by volume. Seven measurements (distributions differentiated by colour) were performed with similar results, three of them at 291 K, two at 298 K and two at 303 K.

neighbours and is without any special biological significance. Furthermore, the four structures with their mutual differences extend our understanding of the influence of glycosylation on the structure and the role of flexibility of the dimer interface in complex formation.

Glycosylation control has been shown by this study to play a very important role in structural investigations of small mammalian proteins. The fully homogeneously glycosylated extracellular domain of human LLT1 clearly gained from the uniformity of the glycosylation pattern. The longer, yet homogeneous, oligosaccharides seem to participate in crystal formation of the LLT1<sub>glyco</sub> form. Their uniform length and type are important. It is also clear that these longer oligosaccharides would disable the crystal contacts observed for deglycosylated LLT1 in the dimeric structures LLT1<sub>D1</sub> and LLT1<sub>D2</sub>. The deglycosylated protein variant enables the formation of contacts in which GlcNAc plays an important role, but longer antennae would disable such tight packing into a crystal. Both forms (uniformly glycosylated and deglycosylated) proved to be important for the structural results, thereby showing the extreme importance of controlled post-translational modifications for detailed molecular studies.

The hydrogen-bonding pattern at the LLT1 dimer interface is not the governing interaction in dimerization and can be significantly weakened, as is the case for the LLT1<sub>glyco</sub> form. Here, the interactions with the neighbouring dimers in the hexamer are strong enough to slightly deform the dimer so that several interface hydrogen bonds near the termini region disappear (the hydrophobic core remains unchanged). The variability of the CTL receptor or ligand dimer interface reported previously, and confirmed by this study, is thus clearly related to the CTL dimer–dimer interaction for the first time. The capacity of these small proteins to form very flexible but still well defined dimers is very likely to form a part of their overall strategy for mediating plastic cell–cell interactions in the immune system.

### Acknowledgements

This study was supported by BIOCEV CZ.1.05/1.1.00/02.0109 from the ERDF, the Ministry of Education, Youth and Sports of the Czech Republic (grants EE2.3.30.0029 and LG14009), the Czech Science Foundation (project P302/11/0855 and 15-15181S), Charles University (GAUK 403211/2010, UNCE 204025/2012 and SVV 260079/2014), the High Education Development Fund (FRVS 669/2013), P-CUBE (EC FP7 project 227764), BioStruct-X (EC FP7 project 283570) and Instruct, part of the European Strategy Forum on Research Infrastructures (ESFRI) supported by national member subscriptions. We thank the staff of beamlines I02 and I04-1 at Diamond Light Source and BM14.1 at BESSY II for technical support.

### References

Aldemir, H., Prod'homme, V., Dumaurier, M.-J., Retiere, C., Poupon, G., Cazareth, J., Bihl, F. & Braud, V. M. (2005). *J. Immunol.* **175**, 7791–7795.

Aust, J. G., Gays, F., Mickiewicz, K. M., Buchanan, E. & Brooks, C. G. (2009). *J. Immunol.* **183**, 106–116.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.

Bartel, Y., Bauer, B. & Steinle, A. (2013). *Front. Immunol.* **4**, 362.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., Bordoli, L. & Schwede, T. (2014). *Nucleic Acids Res.* **42**, W252–W258.

Bláha, J., Páchl, P., Novák, P. & Vaněk, O. (2015). *Protein Expres. Purif.* **109**, 7–13.

Caligiuri, M. A. (2008). *Blood*, **112**, 461–469.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.

Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sust, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J. & Damborsky, J. (2012). *PLoS Comput. Biol.* **8**, e1002708.

Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Mol. Biol.* **4**, 269–275.

Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G. & Baker, N. A. (2007). *Nucleic Acids Res.* **35**, W522–W525.

Durocher, Y., Perret, S. & Kamen, A. (2002). *Nucleic Acids Res.* **30**, E9.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.

Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204–1214.

Fine, J. H., Chen, P., Mesci, A., Allan, D. S., Gasser, S., Raulet, D. H. & Carlyle, J. R. (2010). *Cancer Res.* **70**, 7102–7113.

French, S. & Wilson, K. (1978). *Acta Cryst.* **A34**, 517–525.

Germain, C., Bihl, F., Zahn, S., Poupon, G., Dumaurier, M. J., Rampanarivo, H. H., Padkjaer, S. B., Spee, P. & Braud, V. M. (2010). *J. Biol. Chem.* **285**, 36207–36215.

Germain, C., Meier, A., Jensen, T., Knappnougél, P., Poupon, G., Lazzari, A., Neisig, A., Håkansson, K., Dong, T., Wagtmann, N., Galsgaard, E. D., Spee, P. & Braud, V. M. (2011). *J. Biol. Chem.* **286**, 37964–37975.

Gouet, P., Robert, X. & Courcelle, E. (2003). *Nucleic Acids Res.* **31**, 3320–3323.

Grobárová, V., Benson, V., Rozbeský, D., Novák, P. & Černý, J. (2013). *Immunol. Lett.* **156**, 110–117.

Grueninger-Leitch, F., D'Arcy, A., D'Arcy, B. & Chène, C. (1996). *Protein Sci.* **5**, 2617–2622.

Huarte, E., Cubillos-Ruiz, J. R., Nesbeth, Y. C., Scarlett, U. K., Martinez, D. G., Engle, X. A., Rigby, W. F., Pioli, P. A., Guyre, P. M. & Conejo-Garcia, J. R. (2008). *Blood*, **112**, 1259–1268.

Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.

Kamishikiryo, J., Fukuhara, H., Okabe, Y., Kuroki, K. & Maenaka, K. (2011). *J. Biol. Chem.* **286**, 23823–23830.

Keegan, R. M., Long, F., Fazio, V. J., Winn, M. D., Murshudov, G. N. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 313–323.

Kolenko, P., Rozbeský, D., Vaněk, O., Bezouška, K., Hašek, J. & Dohnálek, J. (2011). *Acta Cryst.* **F67**, 1519–1523.

Kolenko, P., Rozbeský, D., Vaněk, O., Kopecký, V. Jr, Hofbauerová, K., Novák, P., Pompach, P., Hašek, J., Skálová, T., Bezouška, K. & Dohnálek, J. (2011). *J. Struct. Biol.* **175**, 434–441.

Kolenko, P., Skálová, T., Vaněk, O., Štěpánková, A., Dušková, J., Hašek, J., Bezouška, K. & Dohnálek, J. (2009). *Acta Cryst.* **F65**, 1258–1260.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Kveberg, L., Dai, K.-Z., Westgaard, I. H., Daws, M. R., Fossum, S., Naper, C. & Vaage, J. T. (2009). *Eur. J. Immunol.* **39**, 541–551.

Lanier, L. L. (2008). *Nature Immunol.* **9**, 495–502.

Lebedev, A. A. & Isupov, M. N. (2014). *Acta Cryst.* **D70**, 2430–2443.

- Leslie, A. G. W. & Powell, H. R. (2007). *Evolving Methods for Macromolecular Crystallography*, edited by R. J. Read & J. L. Sussman, pp. 41–51. Dordrecht: Springer.
- Li, H., Robertson, A. D. & Jensen, J. H. (2005). *Proteins*, **61**, 704–721.
- Li, Y., Wang, Q., Chen, S., Brown, P. H. & Mariuzza, R. A. (2013). *Proc. Natl Acad. Sci. USA*, **110**, 11505–11510.
- Llera, A. S., Viedma, F., Sánchez-Madrid, F. & Tormo, J. (2001). *J. Biol. Chem.* **276**, 7312–7319.
- Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* **D64**, 125–132.
- Mueller, U., Darowski, N., Fuchs, M. R., Förster, R., Hellmig, M., Paithankar, K. S., Pühringer, S., Steffien, M., Zocher, G. & Weiss, M. S. (2012). *J. Synchrotron Rad.* **19**, 442–449.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Natarajan, K., Sawicki, M. W., Margulies, D. H. & Mariuzza, R. A. (2000). *Biochemistry*, **39**, 14779–14786.
- Ortega, A., Amorós, D. & García de la Torre, J. (2011). *Biophys. J.* **101**, 892–898.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 13419–13424.
- Rosen, D. B., Bettadapura, J., Alsharif, M., Mathew, P. A., Warren, H. S. & Lanier, L. L. (2005). *J. Immunol.* **175**, 7796–7799.
- Rosen, D. B., Cao, W., Avery, D. T., Tangye, S. G., Liu, Y.-J., Houchins, J. P. & Lanier, L. L. (2008). *J. Immunol.* **180**, 6508–6517.
- Roth, P., Mittelbronn, M., Wick, W., Meyermann, R., Tatagiba, M. & Weller, M. (2007). *Cancer Res.* **67**, 3540–3544.
- Satkunanathan, S., Kumar, N., Bajorek, M., Purbhoo, M. A., Culley, F. J. & Lyles, D. (2014). *J. Virol.* **88**, 2366–2373.
- Schuck, P. (2003). *Anal. Biochem.* **320**, 104–124.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol. Syst. Biol.* **7**, 539.
- Skálová, T., Kotýnková, K., Dušková, J., Hašek, J., Koval, T., Kolenko, P., Novák, P., Man, P., Hanč, P., Vaněk, O., Bezouška, K. & Dohnálek, J. (2012). *J. Immunol.* **189**, 4881–4889.
- Sovová, Ž., Kopecký, V. Jr, Pazderka, T., Hofbauerová, K., Rozbeský, D., Vaněk, O., Bezouška, K. & Ettrich, R. (2011). *J. Mol. Model.* **17**, 1353–1370.
- Strong, M., Sawaya, M. R., Wang, S., Phillips, M., Cascio, D. & Eisenberg, D. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 8060–8065.
- Sullivan, L. C., Clements, C. S., Beddoe, T., Johnson, D., Hoare, H. L., Lin, J., Huyton, T., Hopkins, E. J., Reid, H. H., Wilce, M. C. J., Kabat, J., Borrego, F., Coligan, J. E., Rossjohn, J. & Brooks, A. G. (2007). *Immunity*, **27**, 900–911.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Vaněk, O. *et al.* (2008). *FEBS J.* **275**, 5589–5606.
- Vivier, E., Raulet, D. H., Moretta, A., Caligiuri, M. A., Zitvogel, L., Lanier, L. L., Yokoyama, W. M. & Ugolini, S. (2011). *Science*, **331**, 44–49.
- Vivier, E., Tomasello, E., Baratin, M., Walzer, T. & Ugolini, S. (2008). *Nature Immunol.* **9**, 503–510.
- Voigt, S., Mesci, A., Ettinger, J., Fine, J. H., Chen, P., Chou, W. & Carlyle, J. R. (2007). *Immunity*, **26**, 617–627.
- Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.
- Williams, K. J., Wilson, E., Davidson, C. L., Aguilar, O. A., Fu, L., Carlyle, J. R. & Burshtyn, D. N. (2012). *J. Immunol.* **188**, 4980–4991.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yokoyama, W. M. & Plougastel, B. F. M. (2003). *Nature Rev. Immunol.* **3**, 304–316.
- Zelensky, A. N. & Gready, J. E. (2005). *FEBS J.* **272**, 6179–6217.