LETTER

doi:10.1038/nature09448

The structural basis for autonomous dimerization of the pre–T–cell antigen receptor

Siew Siew Pang[1], Richard Berry[1], Zhenjun Chen[2], Lars Kjer-Nielsen[2], Matthew A. Perugini[3], Glenn F. King[4], Christina Wang[1], Sock Hui Chew[1], Nicole L. La Gruta[2], Neal K. Williams[1], Travis Beddoe[1], Tony Tiganis[1], Nathan P. Cowieson[1]†, Dale I. Godfrey[2], Anthony W. Purcell[3], Matthew C. J. Wilce[1], James McCluskey[2] & Jamie Rossjohn[1]

The pre-T-cell antigen receptor (pre-TCR), expressed by immature thymocytes, has a pivotal role in early T-cell development, including TCR β-selection, survival and proliferation of $CD4^-CD8^-$ double-negative thymocytes, and subsequent αβ T-cell lineage differentiation[1–3]. Whereas αβTCR ligation by the peptide-loaded major histocompatibility complex initiates T-cell signalling[4], pre-TCR-induced signalling occurs by means of a ligand-independent dimerization event[5]. The pre-TCR comprises an invariant α-chain (pre-Tα) that pairs with any TCR β-chain (TCRβ) following successful TCR β-gene rearrangement[6]. Here we provide the basis of pre-Tα–TCRβ assembly and pre-TCR dimerization. The pre-Tα chain comprised a single immunoglobulin-like domain that is structurally distinct from the constant (C) domain of the TCR α-chain[7]; nevertheless, the mode of association between pre-Tα and TCRβ mirrored that mediated by the Cα–Cβ domains of the αβTCR. The pre-TCR had a propensity to dimerize in solution, and the molecular envelope of the pre-TCR dimer correlated well with the observed head-to-tail pre-TCR dimer. This mode of pre-TCR dimerization enabled the pre-Tα domain to interact with the variable (V) β domain through residues that are highly conserved across the Vβ and joining (J) β gene families, thus mimicking the interactions at the core of the αβTCR's Vα–Vβ interface. Disruption of this pre-Tα–Vβ dimer interface abrogated pre-TCR dimerization in solution and impaired pre-TCR expression on the cell surface. Accordingly, we provide a mechanism of pre-TCR self-association that allows the pre-Tα chain to simultaneously 'sample' the correct folding of both the V and C domains of any TCR β-chain, regardless of its ultimate specificity, which represents a critical checkpoint in T-cell development. This unusual dual-chaperone-like sensing function of pre-Tα represents a unique mechanism in nature whereby developmental quality control regulates the expression and signalling of an integral membrane receptor complex.

How the extracellular domain of the pre-Tα chain is able to 'report' productive TCR β-gene rearrangement and proper β-chain folding through pairing with any successfully rearranged TCR β-chain, comprising variable V, D and J segments, was unclear. Therefore, we expressed and purified the extracellular domains of pre-Tα–TCRβ and determined its crystal structure (Supplementary Table 1). The pre-Tα domain is folded into an immunoglobulin domain (Fig. 1a), best described as a C-type topology, and comprises a compact β-sandwich of two antiparallel β-sheets tethered by the canonical intramolecular disulphide bond, with the amino and carboxy termini at opposite ends of the molecule (Fig. 1a). The pre-Tα domain was more structurally similar to the C domain of an antibody light chain (IgλC; 34% sequence similarity; root mean squared deviation of 5.1 Å for 73 Cα atoms) in comparison with the Cα domain (26% sequence similarity; root mean squared deviation 12.6 Å for 57 Cα atoms) (Supplementary Fig. 1a). Indeed, the pre-Tα fold contrasted with that of the atypical immunoglobulin fold of the αβTCR's Cα domain[7] (Fig. 1b), which is characterized by one β-sheet that abuts a number of loosely packed random coils. Moreover, the pre-Tα domain was notably elongated (42 Å) in comparison with its Cα counterpart (33 Å), an elongation that is attributable to longer A, B and E β-strands, and the E–F and A–B loops (Fig. 1). Although the 'top' of the pre-Tα domain is unusually hydrophobic, flat and featureless (Supplementary Fig. 1b, c), a large cluster of hydrophobic residues is present on one face of the β-sheet (Supplementary Fig. 1c) and a Trp residue is prominently exposed on the other face (Supplementary Fig. 1b,c). These are features that typically mediate protein–protein interactions, and indeed they were observed to have a role in mediating pre-TCR assembly (see below).

Within the pre-Tα–TCRβ complex, the conformation of the TCR β-chain closely matched that observed in the parental LC13 αβTCR[8], revealing that the lack of a Vα domain in the pre-TCR does not result in the crumpling of the Vβ domain (Fig. 2a, b). Notably, this asymmetrical shape leads to a large, surface-exposed hydrophobic patch on the unpaired Vβ domain within the pre-Tα–TCRβ complex, leaving Tyr 35 and Phe 108 solvent exposed (Fig. 2a). Within the pre-TCR, the pre-Tα domain interacts exclusively with the Cβ domain in a clasp-like interaction (Fig. 2a). Collectively, this forms a continuous and extensive interface with a buried surface area (BSA) of ~1,770 $Å^2$, albeit with poor

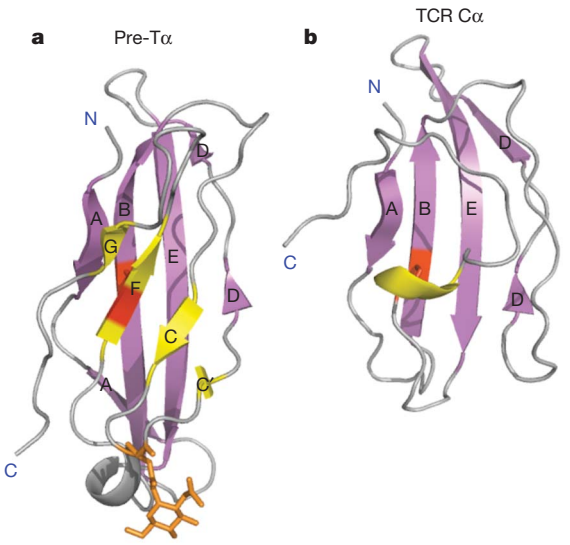

Figure 1 | The pre-Tα structure. Immunoglobulin folds of pre-Tα (**a**) and the Cα domain (**b**), with disulphide bond in red.

[1]The Protein Crystallography Unit, Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia. [2]Department of Microbiology & Immunology, University of Melbourne, Parkville, Victoria 3010, Australia. [3]Department of Biochemistry and Molecular Biology and Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria 3010, Australia. [4]Division of Chemistry & Structural Biology, Institute for Molecular Bioscience, The University of Queensland, St Lucia, Queensland 4072, Australia. †Present address: The Australian Synchrotron, Blackburn Road, Monash, Clayton, Victoria 3168, Australia.

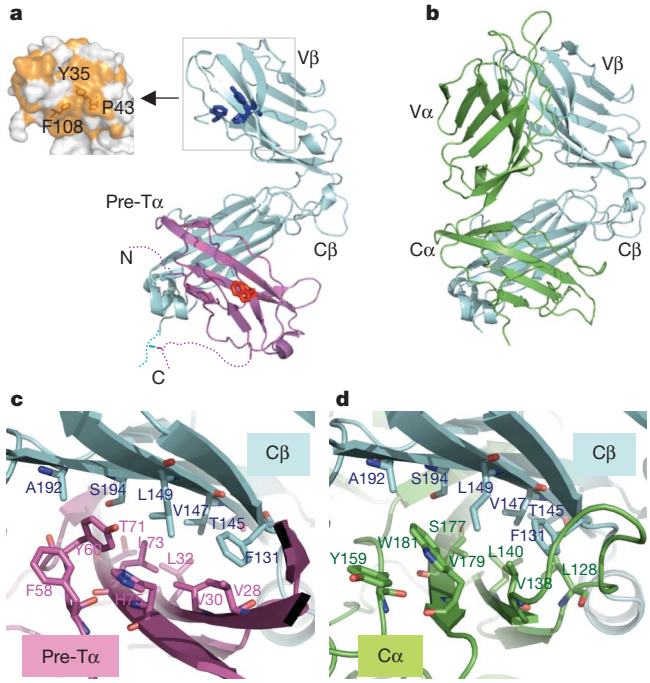

Figure 2 | Comparison between the pre-TCR and αβTCR structures. **a**, Pre-Tα (magenta) interacts with the Cβ domain. The hydrophobic patch is shown as a blue stick model (also see inset); Trp 46, red. Dotted lines represent the N and C termini, which are not visible in the crystal structure. **b**, αβTCR (Protein Data Bank ID, 1KGC[9]) contains two chains (α, green; β, cyan), with Vα interacting with Vβ and Cα interacting with Cβ. **c**, **d**, The contact interfaces between pre-Tα (magenta) and Cβ (cyan) of pre-TCR (**c**) and between Cα (green) and Cβ (cyan) of αβTCR (**d**). The contact residues are shown as stick models.

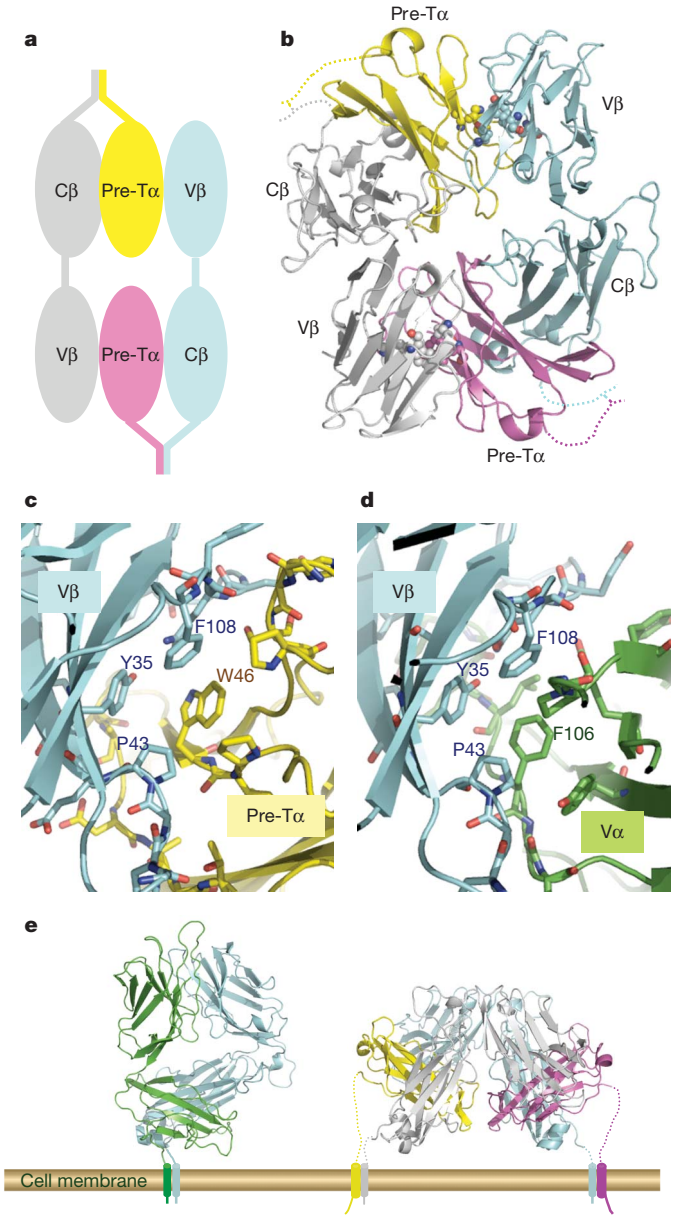

Figure 3 | The pre-TCR dimer. **a**, Pre-TCR (pre-Tα, magenta; β-chain, cyan) interacts with a second receptor (pre-Tα, yellow; β-chain, white) in a head-to-tail arrangement. **b**, Dimerization is mediated mainly between the pre-Tα and the unpaired Vβ domains. Some of the residues involved at these interfaces are shown as space-filling models. **c**, **d**, Contact interface between pre-Tα (yellow) and Vβ (cyan) of pre-TCR (**c**) and between Vα (green) and Vβ (cyan) of αβTCR (**d**). The highly conserved interacting residues are labelled and shown as stick models. **e**, Cartoon showing orientation of the TCRs on the cell surface: left, mature αβTCR; right, pre-TCR.

shape complementarity, of 0.57 (Fig. 2c and Supplementary Table 2). The core of the pre-Tα–Cβ interface is dominated by van der Waals contacts, including interactions mediated by a number of small polar or aliphatic residues along the surface of the β-sheets and a cluster of aromatic residues (Fig. 2c and Supplementary Table 2). The location of the pre-Tα–Cβ interface coincides with that of the Cα–Cβ interface, although the latter interface is much more extensive[9] (BSA ≈ 2,500 Å^2; Fig. 2d). The larger Cα–Cβ interface arises mainly from the longer D–E loop of the Cα domain. Nevertheless, core interactions at the Cα–Cβ interface are mimicked at the pre-Tα–Cβ interface, involving 17 common Cβ contact residues that include a series of interlocking small hydrophobic residues between the β-sheets and a small aromatic cluster (Fig. 2d and Supplementary Table 2), which suggests that hydrophobic forces drive pre-Tα–TCRβ assembly[9]. The larger BSA and greater shape complementarity (0.67) at the Cα–Cβ interface is consistent with the ability of the TCR α-chain to compete off the pre-Tα domain effectively during αβTCR development.

The pre-TCR is known to self-associate on the cell surface of CD4$^-$CD8$^-$ double-negative thymocytes, and spontaneously dimerizes *in vitro* and *in vivo*, with the dimerization being mediated by residues within the pre-Tα domain[5,10]. In the absence of any structural data, this has led to speculative models of the pre-TCR dimeric assembly. Within the crystal lattice, two potential pre-TCR dimers were observed. The first pre-TCR dimer involved a side-by-side arrangement in which the total BSA at the interface was small (BSA ≈ 1,380 Å^2), and involved contacts that were mediated mainly through polar residues (Supplementary Fig. 2). This side-by-side pre-TCR dimer would sterically block any interactions the Cβ domain is reported to make with the CD3 complex[11]. Furthermore, the shape of this side-by-side pre-TCR dimer did not match what was observed in solution (see below). Accordingly, the side-by-side pre-TCR dimer was considered to be attributable solely to the crystal contacts and was discounted from further analyses. The second pre-TCR dimer (Fig. 3) observed in the crystal lattice was markedly more biophysically and biologically convincing, and involved a head-to-tail dimeric arrangement of the pre-TCR heterodimers. Here the pre-Tα domain of one protomer was sandwiched between the Cβ and Vβ domains (Fig. 3a, b). The hydrophobic and featureless 'top' of the pre-Tα domain was critical for this head-to-tail dimeric arrangement, as large, bulky and charged residues would have prevented this mode of dimerization. The pre-TCR head-to-tail dimerization creates two new interfaces, each involving the C′CGF β-sheet of pre-Tα from one receptor interacting with the β-sheet of the unpaired Vβ domain from the other (Fig. 3b and Supplementary Table 3). This arrangement occludes the otherwise surface-exposed hydrophobic patch on the Vβ domain (Fig. 2a) and provides immediate insight into how and why the

pre-TCR autonomously dimerizes in a *cis*-configuration on the cell surface (Fig. 3e). By pairing with the Vβ domain, the location of the pre-Tα domain mimics the positioning of the Vα domain at the Vα-Vβ interface, although different faces of the domains are used (Fig. 2b). The total BSA at the head-to-tail pre-TCR dimer interface was large (BSA ≈ 3,000 Å^2) and comprised a large number of van der Waals interactions (Supplementary Table 3), as is typically observed within oligomeric protein–protein complexes. Unlike in previous theories of pre-TCR dimerization[5], charged residues within the pre-Tα domain did not have a role in mediating pre-TCR dimerization (Supplementary Fig. 3), and thus may be involved in other ways, such as in pre-TCR–CD3 assembly. Consistent with this, some of the mutations known to affect pre-TCR function reside within the A–B loop of the pre-Tα domain (Asp 22 and Lys 24), a loop that is implicated in αβTCR–CD3 interactions[12]. Importantly, the pre-Tα–Vβ interface principally involved residues located on the C′CGF-sheet of the pre-Tα domain (Fig. 3c), a region that is structurally dissimilar from the Cα domain, thereby providing an explanation of why the corresponding Cα domain cannot act as a suitable surrogate for the pre-Tα domain in mediating this head-to-tail dimeric assembly and, hence, pre-TCR signalling. Within this interface, which covers ~1,260 Å^2, the invariant pre-Tα Trp 46 is positioned at the core (Fig. 3c and Supplementary Table 3). This ball-and-socket interaction mirrored what was observed at the LC13 Vα–Vβ interface (BSA ≈ 1,490 Å^2), in which Phe 106 from Vα is the corresponding central residue; furthermore, there were 12 common Vβ contact points between pre-Tα–Vβ and Vα–Vβ (Fig. 3d and Supplementary Table 3). Notably, pre-Tα makes contact with residues encoded by the Vβ domain (Tyr 35 and Pro 43) and the Jβ region (Phe 108), thus providing a mechanism that enables the pre-Tα domain to 'sense' the correct folding of the various components derived from disparate gene segments that create the TCR β-chain. Within the TCR β-locus, there are 42 Vβ gene segments, but of these two Vβ contact points, position 43 is invariably occupied by a short hydrophobic residue (Pro, Leu, Ile or Val), and Tyr 35 is invariant. Furthermore, Phe 108 is almost invariant within the 12 Jβ gene segments (Supplementary Fig. 4). This conserved constellation of interacting residues provides a framework for understanding how pre-Tα can indiscriminately pair with any successfully rearranged TCR β-chain.

Next, to further verify that the pre-TCR could dimerize *in vitro*, we investigated the behaviour of the pre-TCR in solution. As judged by sedimentation velocity analytical ultracentrifugation (SV-AUC) and small-angle X-ray scattering (SAXS), the pre-TCR displayed the hallmarks of a weak self-associating system, typical of many immunological receptors[13]. Specifically, using SAXS analysis, we observed the pre-TCR to undergo a concentration-dependent increase in molecular mass, radius of gyration (R_g) and maximal particle dimension (Supplementary Table 4). These effects cannot be accounted for by the presence of non-specific protein aggregation, as judged by the linearity of SAXS Guinier plots (Supplementary Fig. 5). Furthermore, using SV-AUC, we see two distinct species (Fig. 4a). On the basis of hydrodynamic calculations, these species correspond to pre-Tα–TCRβ (apparent sedimentation coefficient, $s_{app} = 3.1$ S; frictional ratio, $f/f_0 = 1.44$) and the pre-TCR dimer ($s_{app} = 5.0$ S, $f/f_0 = 1.37$) (Fig. 4b and Supplementary Table 5). Consistent with the AUC results, the SAXS measurements show an increase in average particle size with increasing protein concentration, consistent with a monomer–dimer equilibrium (Supplementary Fig. 6a). To gain an insight into the shape of the pre-TCR in solution, low-resolution models were restored from the SAXS scattering curves *ab initio* using the DAMMIF[14] and GASBOR[15] programs (Supplementary Fig. 6a). As the molecular weight determined by SAXS suggests that the pre-TCR is fully dimeric at 15 mg ml^{-1}, we fitted the pre-TCR dimer into this SAXS model. In contrast to the side-by-side pre-TCR dimer (Supplementary Fig. 6b), the head-to-tail pre-TCR dimer convincingly fitted the central globular region of the SAXS model (Fig. 4f and Supplementary Fig. 6c). Furthermore, in the head-to-tail pre-TCR dimer, the regions that

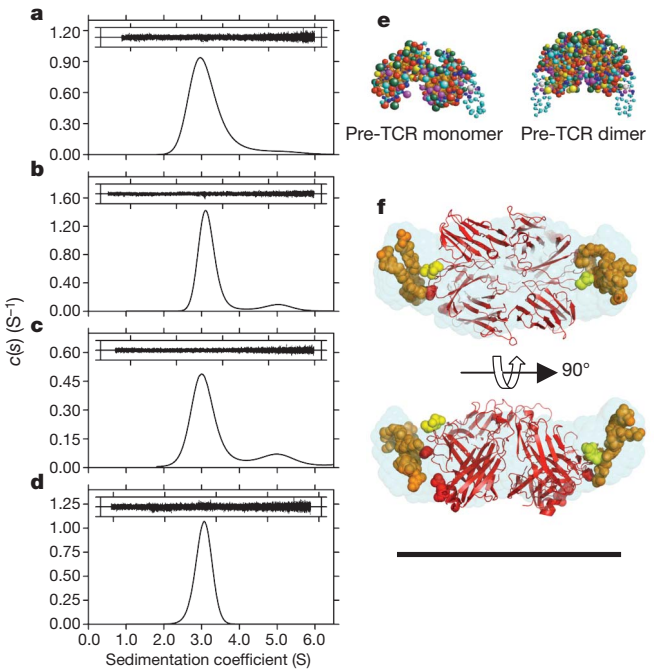

Figure 4 | Pre-TCR dimerization in solution. **a–d**, Continuous size distributions, determined by SV-AUC, of pre-TCR at 1 mg ml^{-1} (**a**), 3 mg ml^{-1} (**b**) and 8.7 mg ml^{-1} (**c**), and of W46αR pre-TCR mutant at 1.4 mg ml^{-1} (**d**). Insets show the corresponding residuals. **e**, Bead models generated from the pre-TCR structure for the monomer (left) and the head-to-tail dimer (right). **f**, Overlay of the pre-TCR head-to-tail dimer (red ribbon) and the DAMMIF SAXS model (cyan spheres), shown in two orientations. Red dots indicate C-terminal residues of the pre-Tα and LC13 β-chains, orange dots indicate modelled C-terminal tails and yellow dots indicate N-linked glycosylation sites. Scale bar, 100 nm.

are missing from the crystal structure (principally the C-terminal stalks of both pre-Tα and the TCR β-chains) are present at opposing poles, where they are well positioned to occupy the areas left vacant by the crystal structure. Thus, the head-to-tail pre-TCR dimer observed in the crystal structure correlated well with the differing solution-based measurements. This *cis*-dimeric arrangement suggests that the pre-TCR sits 'flat', almost parallel to the membrane surface, in contrast to the αβTCR, which is usually depicted in an upright orientation on the membrane (Fig. 3e). Consistent with this, the relative dependency of the various CD3 subunits within the pre-TCR complex may differ from that within the αβTCR–CD3 complex[16].

Next we used three distinct approaches to further validate the pre-TCR dimer observed in the crystal structure and in solution. First, we used crosslinking/mass spectrometry studies of recombinant pre-TCR to directly assess the configuration of the dimer (Supplementary Fig. 7 and Supplementary Table 6; see Methods). Notably, the only crosslinked species between the pre-Tα chain and the TCR β-chain that we observed was between the N-terminal sequence (head) of the Vβ domain and the C-terminal sequence (tail) of the pre-Tα chain (Supplementary Fig. 7 and Supplementary Table 6). This crosslinked species is incompatible with the pre-Tα–TCRβ and the side-by-side pre-TCR dimer, but is fully consistent with a head-to-tail pre-TCR dimer configuration. Second, we generated a W46αR mutant in the pre-Tα chain and assessed its impact on pre-TCR dimerization as judged by SV-AUC. In contrast to the wild-type (WT) pre-TCR, sedimentation velocity analysis of the W46αR mutant shows that it is monomeric in solution ($s_{app} = 3.1$ S) with no evidence of dimerization (Fig. 4). Third, to assess the pre-TCR dimer conformation on the cell surface, we transfected WT and mutant pre-TCRs into the TCR αβ-negative, CD3-positive SKW3 T-cell line and visualized cell surface and intracellular pre-TCR localization (Fig. 5). As a control, we also transfected

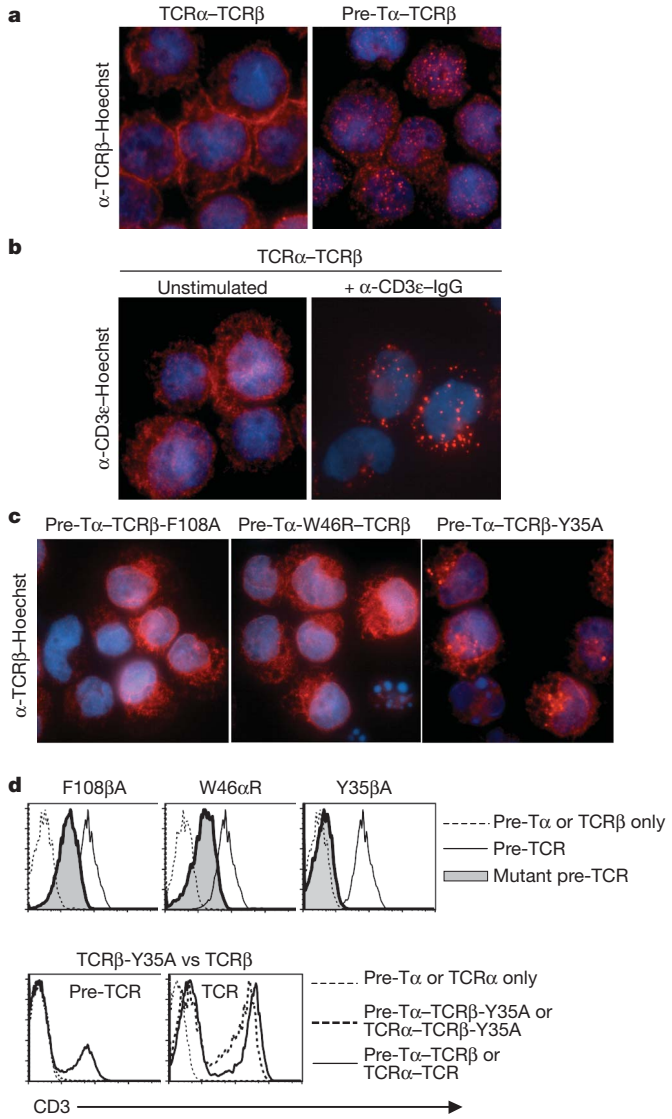

Figure 5 | Expression of pre-TCR mutants. **a**, TCRβ staining of SKW3 T cells expressing TCRα–TCRβ or pre-Tα–TCRβ. Hoechst, Hoechst stain. **b**, CD3ε staining of unstimulated and activated SKW3 T cells expressing TCRα–TCRβ. IgG, immunoglobulin G. **c**, TCRβ staining of T cells expressing the indicated pre-Tα and TCRβ mutations. **d**, Flow staining (CD3ε) of T cells expressing pre-Tα and TCRVβ mutations. The top row shows negative-control T cells (either pre-Tα or LC13β alone; thin dashed lines), positive-control pre-Tα–LC13β transfectants (solid lines) and SKW3.pre-Tα+LC13β mutants (grey shading). The bottom row shows unsorted T cells expressing either pre-Tα–LC13β-Y35A, pre-Tα alone or pre-Tα–LC13β (left), and unsorted SKW3 cells expressing LC13α–LC13β-Y35A, LC13α–LC13β or TCRα alone (negative control; right). See also Supplementary Fig. 8.

the parental LC13 αβTCR into the SKW3 cell line. We demonstrated, using immunofluorescence, that transfection of the WT pre-TCR resulted in a punctate expression pattern, typical of pre-TCR clustering on the cell surface and consequent endocytosis, which is in direct contrast to the linear cell surface expression pattern of the mature αβTCR in unstimulated cells (Fig. 5a) and the large puncta resulting from cross-linking the αβTCR on the cell surface with α-CD3ε (Fig. 5b). Next we examined the impact of three pre-TCR mutants, Y35βA, F108βA and W46αR, on cell surface expression and intracellular localization using flow cytometric analysis and immunofluorescence microscopy. The F108βA and W46αR pre-TCR mutants were expressed at lower levels on the cell surface and led to a greater retention of these mutants in an intracellular compartment typical of the endoplasmic reticulum (Fig. 5c, d). Furthermore, the Y35βA TCR–Vβ mutant completely abrogated the cell surface expression of the pre-TCR (Fig. 5d) and resulted in accumulation of the expressed pre-TCR Y35βA mutant in perinuclear puncta reminiscent of the Golgi complex and the endoplasmic reticulum, consistent with defective trafficking to the cell surface (Fig. 5c). However, the Y35βA mutant did not affect LC13 αβTCR cell surface expression (Fig. 5d), indicating that this mutant does not affect the fold of the TCR β-chain *per se*. This finding highlights the fact that pre-TCR cell surface expression is finely tuned to pre-Tα–Vβ interactions. Collectively, the crosslinking experiments in solution, coupled with the effect of the mutants on pre-TCR dimerization in solution and on the cell surface, are fully consistent with the head-to-tail pre-TCR dimer configuration.

In vertebrates there are three classes of TCR, namely αβTCR, γδTCR and the pre-TCR. Although structural data exists on the architecture of αβTCR[7] and γδTCRs[17], information is lacking on the pre-TCR. Furthermore, whereas αβTCR and γδTCR interact with defined ligands, the pre-TCR is unique in that a ligand-independent mechanism underpins its role in T-cell development. Our findings formally demonstrate that the pre-Tα chain functions much more than a simple surrogate TCR α-chain, whereby the pre-Tα domain simultaneously binds to the Cβ and Vβ domains within the head-to-tail pre-TCR dimer. The role of the pre-Tα domain is to act as a checkpoint that assesses correct rearrangement and expression of the TCR β-chain on the cell surface. The observed mode of dimerization can only occur when both the variable and constant domains of the TCR β-chain are correctly folded; this is a previously unrecognized determinant for β-selection. Thus, the pre-Tα has a dual chaperone-sensing role that is synchronously linked to dimerization and pre-TCR signal transduction. Many proteins require chaperones for correct folding, and subsequent function, with intricate cellular machines typically being required to aid in this process. The pre-B-cell receptor[18] has a number of components that allow productive expression of the B-cell receptor in a manner that is quite distinct from the individual pre-TCR. The evolutionarily conserved pre-Tα has achieved its role by adopting dual functionality and, as such, represents a previously undescribed mechanism in nature that ensures the integrity of protein assembly and function.

METHODS SUMMARY

Pre-TCR cloning, expression and purification. We cloned the ectodomains of the human pre-Tα and TCR LC13β into the pFastBac Dual vector (Invitrogen). The pre-TCR was expressed, purified and crystallized, and its structure determined. Further details are provided in Methods.

SAXS. SAXS data were collected at the Australian Synchrotron, Melbourne, and the Advanced Light Source, Berkeley. We assessed data quality using Guinier analysis, and determined R_g and the maximal particle dimension, D_{max}, using GNOM[19]. *Ab initio* models were generated using DAMMIF and GASBOR and were overlaid with high-resolution models using PYMOL (http://www.pymol.org/).

Mass spectrometry and crosslinking studies. These studies were performed on recombinant pre-TCR as described in Methods.

Functional studies. We assessed pre-TCR and mutants of pre-TCR, as well as αβTCR and mutants of αβTCR, for cell surface expression as described in Methods. T cells were processed for epifluorescence microscopy as described in Methods.

AUC. AUC experiments on WT and mutant pre-TCR were conducted using a Beckman model XL-I analytical ultracentrifuge at a temperature of 20 °C using standard protocols and analyses. Bead models representing pre-TCR were generated using SOMO[20]. See Methods for more details.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 April; accepted 23 August 2010.
Published online XX 2010.

1. von Boehmer, H. Unique features of the pre-T-cell receptor α-chain: not just a surrogate. *Nature Rev. Immunol.* **5,** 571–577 (2005).
2. Saint-Ruf, C. *et al.* Different initiation of pre-TCR and γδTCR signalling. *Nature* **406,** 524–527 (2000).

3. Borowski, C., Li, X., Aifantis, I., Gounari, F. & von Boehmer, H. Pre-TCRα and TCRα are not interchangeable partners of TCRβ during T lymphocyte development. *J. Exp. Med.* **199,** 607–615 (2004).
4. Zinkernagel, R. M. & Doherty, P. C. Restriction of *in vitro* T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature* **248,** 701–702 (1974).
5. Yamasaki, S. *et al.* Mechanistic basis of pre-T cell receptor-mediated autonomous signaling critical for thymocyte development. *Nature Immunol.* **7,** 67–75 (2005).
6. Saint-Ruf, C. Analysis and expression of a cloned pre-T cell receptor gene. *Science* **266,** 1208–1212 (1994).
7. Garcia, K. C. *et al.* An αβ T cell receptor structure at 2.5 Å and its orientation in the TCR-MHC complex. *Science* **274,** 209–219 (1996).
8. Kjer-Nielsen, L. *et al.* A structural basis for the selection of dominant αβ T cell receptors in antiviral immunity. *Immunity* **18,** 53–64 (2003).
9. Kjer-Nielsen, L. *et al.* The 1.5 Å crystal structure of a highly selected antiviral T cell receptor provides evidence for a structural basis of immunodominance. *Structure* **10,** 1521–1532 (2002).
10. Yamasaki, S. & Saito, T. Molecular basis for pre-TCR-mediated autonomous signaling. *Trends Immunol.* **28,** 39–43 (2007).
11. Kuhns, M. S. & Davis, M. M. Disruption of extracellular interactions impairs T cell receptor-CD3 complex stability and signaling. *Immunity* **26,** 357–369 (2007).
12. Beddoe, T. *et al.* Antigen ligation triggers a conformational change within the constant domain of the αβ T cell receptor. *Immunity* **30,** 777–788 (2009).
13. van der Merwe, P. A. & Davis, S. J. Molecular interactions mediating T cell antigen recognition. *Annu. Rev. Immunol.* **21,** 659–684 (2003).
14. Franke, D. & Svergun, D. I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42,** 342–346 (2009).
15. Svergun, D. I., Petoukhov, M. V. & Koch, M. H. Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.* **80,** 2946–2953 (2001).
16. Dave, V. P. Hierarchical role of CD3 chains in thymocyte development. *Immunol. Rev.* **232,** 22–33 (2009).
17. Allison, T. J., Winter, C. C., Fournie, J. J., Bonneville, M. & Garboczi, D. N. Structure of a human γδ T-cell antigen receptor. *Nature* **411,** 820–824 (2001).
18. Melchers, F. The pre-B-cell receptor: selector of fitting immunoglobulin heavy chains for the B-cell repertoire. *Nature Rev. Immunol.* **5,** 578–584 (2005).
19. Semenyuk, A. V. & Svergun, D. I. GNOM - a program package for small-angle scattering data processing. *J. Appl. Crystallogr.* **24,** 537–540 (1991).
20. Brookes, E., Demeler, B., Rosano, C. & Rocco, M. The implementation of SOMO (SOlution MOdeller) in the UltraScan analytical ultracentrifugation data analysis suite: enhanced capabilities allow the reliable hydrodynamic modeling of virtually any kind of biomacromolecule. *Eur. Biophys. J.* **39,** 423–435 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Turner for discussions; S. Ramarathinam for technical assistance; the staff at the Australian synchrotron (MX and SAXS/WAXS beamlines) and Berkeley Advanced Light Source (SIBYLS beamline) for assistance with data collection; and T. Caradoc-Davies for advice on processing of merohedral twinned X-ray data. This research was supported by grants from the Australian Research Council and the National Health and Medical Research Council of Australia. A.W.P., T.T. and M.C.J.W. are supported by NHMRC Senior Research Fellowships, and D.I.G. is supported by an NHMRC Principal Research Fellowship. M.A.P. is supported by an ARC Future Fellowship and J.R. is supported by an ARC Federation Fellowship.

Author Contributions S.S.P. solved the structure, undertook analysis, performed experiments and contributed to manuscript preparation. J.M., D.I.G. and J.R. contributed to the design and interpretation of experiments, project management and the writing of the manuscript. R.B., Z.C., L.K.-N., M.A.P., G.F.K., N.L.L.G., T.T. and A.W.P. performed experiments and contributed to the writing of the manuscript; C.W., S.H.C., N.K.W., T.B., N.P.C. and M.C.J.W. performed experiments. J.M. and J.R. were the joint senior authors—they co-led the investigation, devised the project, analysed the data and wrote the manuscript.

Author Information The atomic coordinates and structure factors of pre-TCR have been deposited in the Protein Data Bank under the accession code 3OF6. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.M. (jamesm1@unimelb.edu.au) or J.R. (jamie.rossjohn@monash.edu).

METHODS

Pre-TCR cloning, expression and purification. Genes encoding the ectodomains (including the interchain cysteine) of the human pre-Tα and TCR LC13β were cloned into the pFastBac Dual vector (Invitrogen). The genes were cloned downstream of the baculovirus GP67 secretory signal sequence and were under the control of different promoters for co-expression. The C terminus of pre-Tα was modified to include an enterokinase cleavage site, a Jun zipper, a BirA tag and a hexahistidine (6xHis) tag, and the LC13 β-chain has an enterokinase site and a Fos zipper. The pre-TCRs were expressed and secreted by High Five insect cells (Invitrogen) as a covalently linked dimer. At the end of expression, the insect cell supernatant was collected and diafiltrated with TBS by tangential flow filtration. The hexahistidine-tagged proteins were purified using Ni Sepharose High Performance resin (GE Healthcare) and by gel filtration (HiLoad 16/60 Superdex 200, GE Healthcare). For protein crystallization, the C-terminal tag regions were removed by enterokinase digestion (GenScript). Final purification and buffer exchange into 20 mM Tris-HCl, (pH 8.0) and 0.15 M NaCl was carried out by gel filtration. The pooled protein was concentrated to ~10 mg ml^{-1}, snapped frozen and stored at −80 °C in small aliquots. The typical protein yield of the cleaved pre-TCR is about 1–3 mg per litre of expression culture.

Crystallization and X-ray data collection. Nanoscale crystallization trials of pre-TCR were carried out using sitting-drop vapour diffusion at room temperature (19 °C) and at 4 °C. A few hits were identified and reproduced manually using hanging-drop vapour diffusion. The optimized crystallization conditions were produced by mixing equal volumes of pre-TCR (6 mg ml^{-1}) with the reservoir buffer (0.1 M sodium cacodylate (pH 7.0), 0.1 M calcium acetate and 13–16% PEG 1500). Rod-shaped pre-TCR crystals appeared quickly after overnight incubation at 4 °C and continued to grow in size for the next few days. For cryoprotection, the rod-shaped crystals were soaked in the reservoir buffer plus 15% glycerol overnight at 4 °C. The soaked crystals were then transferred into reservoir buffer containing 25% glycerol and flash-frozen in liquid nitrogen. X-ray data to 2.8 Å were collected using the Micro Crystallography (MX2) beamline at the Australian Synchrotron. The diffraction data were processed and analysed using MOSFILM 6.2.6 and SCALA (CCP4 suite)[21]. A summary of data collection statistics is provided in Supplementary Table 1.

Structure determination and refinement. Initial analysis of the pre-TCR crystals indicated that the X-ray data could be scaled into any of the primitive hexagonal space groups, suggesting crystal twinning. Further examination using The Merohedral Crystal Twinning Server (http://nihserver.mbi.ucla.edu/Twinning/), CCP4 programs and PHENIX.XTRIAGE[22] confirmed perfect crystal merohedral twinning. The structure was solved by molecular replacement (PHASER from the CCP4 suite) using the β-chain of LC13 TCR (PDB ID, 1KGC) as the search model. Solutions were found in a number of space groups (six molecules for $P3_2$, three molecules for $P3_221$ or $P3_212$, and one molecule for $P6_222$). Each of the solutions was examined and refined using PHENIX[23], taking into consideration the twinning factor and operator (0.516, $-k$, $h+k$, l; k, h and l are Miller indices) After initial rounds of refinement, the $P3_221$ solution gave a reasonable R_{factor} and an unbiased density of the missing pre-Tα chains was observed, suggesting the correct solution. To avoid bias due to the presence of data twinning, the R_{free} set was assigned in PHENIX.REFINE. Further restrained refinements were performed using PHENIX.REFINE after initial rounds of rigid-body refinement and stimulated annealing. Each round of refinement was interspersed with electron density inspection and manual model building with COOT[24]. The progress of refinement was monitored by R_{free} value and the final refinement statistics for pre-TCR are given in Supplementary Table 1. There were three pre-TCR complexes in the asymmetric unit, all of which were highly similar to each other, and structural analyses were thus confined to one pre-TCR complex. The final model comprises residues 7–110 from the pre-Tα chain and residues 2–245 from the TCR β-chain. Disordered regions include six residues from the N terminus and nine from the C terminus of pre-Tα, and one N- and three C-terminal residues from TCRβ, respectively. The missing regions include the intermolecular disulphide bond at the flexible C-terminal end of the polypeptides.

SAXS. SAXS data were collected at the Australian Synchrotron using a 1M Pilatus detector. Buffers/samples were loaded into 1-mm quartz capillaries and continuously flowed through the beam during data collection. Multiple 1-s exposures were collected and compared to control for radiation damage. Data were collected using a single camera length with a 1.5-m sample-to-detector distance to cover a momentum transfer interval of $0.0056\ Å^{-1} < q < 0.3\ Å^{-1}$. The modulus of the momentum transfer is defined as $q = 4\pi\sin(\theta/\lambda)$, where 2θ is the scattering angle and λ is the wavelength. Scattering images were radially averaged and blank subtracted using SAXS15ID software (Australian Synchrotron). Additional SAXS data were collected at the Advanced Light Source (ALS), Berkeley, using a MAR165CCD detector that was 1.5 m from the sample, resulting in a q range of 0.01–0.3 Å^{-1}. Circular integration and normalization was performed using in-house software. Molecular mass estimates were obtained by normalizing scattering to BSA (ALS) or lysozyme (Australian Synchrotron). Data quality was assessed on the basis of the linearity of Guinier plots, and R_g and the pairwise intraparticle distance distribution function were determined using GNOM[19]. *Ab initio* models were generated using DAMMIF[14] and GASBOR[15] using *P*1 symmetry in all cases except for the 15 mg ml^{-1} concentration that corresponds to pre-TCR dimer (*P*2 symmetry). Ten independent DAMMIF/GASBOR runs were aligned, combined and filtered to generate a final model that retained the most consistent features using the DAMMAVER package. The normalized spatial discrepancies between individual 15 mg ml^{-1} models were 1.17–1.22 (GASBOR) and 0.61–0.66 (DAMMIF). High-resolution models of pre-TCR were fitted within *ab initio* models using PYMOL. Bead models (Fig. 4) generated from the pre-TCR structure for the monomer and head-to-tail pre-TCR dimer include C termini that were manually generated in PYMOL.

AUC. AUC experiments were conducted using a Beckman model XL-I analytical ultracentrifuge at a temperature of 20 °C. Samples (380 µl) of wild-type pre-Tα–LC13β solubilized in 20 mM Tris, 100 mM NaCl and 50 mM Ca acetate (pH 7.1), W46R-mutant pre-Tα–LC13β solubilized in 20 mM Tris, 150 mM NaCl and 1 mM EDTA (pH 7.0), and reference solutions (400 µl) were loaded into a conventional double-sector quartz cell and mounted in a Beckman four-hole An-60 Ti rotor. Intensity data were collected in continuous mode at a single wavelength between 293 and 306 nm using a rotor speed of 40,000 r.p.m. and a step size of 0.003 cm without averaging. Frictional ratios (f/f_0) were calculated with SEDNTERP[25] using the measured sedimentation coefficient of the given species, the molecular mass calculated from the amino-acid sequence (41,800-Da pre-TCR and 83,600-Da pre-TCR dimer), a solvent density of 1.005 g ml^{-1}, a viscosity of 1.021 cP and a partial specific volume of 0.722 ml g^{-1}. Sedimentation velocity intensity data at multiple time points were converted to absorbance data and then fitted to a continuous size distribution model using the program SEDFIT[26] with a resolution of 200 species ranging from 1.0 S (or 10 kDa) to 7.2 S (or 110 kDa) and a *P* value of 0.95.

Mass spectrometry and crosslinking studies. A 2 mg ml^{-1} solution of recombinant pre-TCR was buffer exchanged into PBS and treated with bis(sulfosuccinimidyl) suberate (BS3) crosslinker (5 mM final concentration) for 30 min on ice, and the reaction was quenched with the addition of 200 mM Tris (pH 8.0). The crosslinked material was resolved on a one-dimensional SDS–PAGE gel and the band corresponding to the dimer excised and an in-gel tryptic digest performed. Crosslinked peptides were examined by LC-MS/MS using an AB SCIEX Q-STAR ELITE mass spectrometer and X-QUEST software[27].

Cloning and expression of pre-Tα, LC13α and LC13β. pGEM–LC13α and pMIG–LC13β constructs were generated as previously described[12]. Pre-TCR cDNA was reverse-transcribed from isolated MOLT RNA using Superscript II (Invitrogen). PCR products encoding the pre-Tα gene were cloned into the pGEM–T Easy vector (Promega) and verified by sequencing before being transferred into the retroviral expression vector pMIG. Primers used were as follows: pre-Tα-F, 5′-CGGAATTCACGCGTATGGCCGGTACATGGCTGCT-3′; pre-Tα-R, 5′-CGCTCGAG*CACAAAGTG*TCAGGCAGCTCCAGCCTGCAGAGG-3′. Mutants were created with pGEM–pre-Tα or pGEM–LC13β plasmid DNA as template by site-directed mutagenesis (QuikChange; Stratagene). Primers used are as follows with underlined letters indicating codons mutated: pre-Tα-W46R-F, 5′-cttgacagccccatcCGGttctcagccggcaat-3′; pre-Tα-W46R-R, 5′-attgccggctgagaaCCGgatggggctgtcaag-3′; LC13β-Y35A-F, 5′-CATGTATCCCTTTTTTGGGCCCAACAGGCCCTGGGG-3′; LC13β-Y35A-R, 5′-CCCCAGGGCCTGTTGGGCCCAAAAAAGGGATACATG-3′; LC13β-F108A-F, 5′-GCCTACGAGCAGTACGCCGGGCCGGGCACCAGG-3′; LC13β-F108A-R, 5′-CCTGGTGCCCGGCCCGGCGTACTGCTCGTAGGC-3′.

Functional assays. SKW3 is a human T-cell leukaemia line that is $CD4^+CD8^+$ double positive and $TCRa^-b^-$, $TCRg^-d^-$ deficient[28,29]. Single TCR gene-transfectants SKW3.pre-Tα, SKW3.LC13α and SKW3.LC13β were generated by retroviral transduction of the TCR α- and β-chains into SKW3 cells as described previously[30]. Transfected cells (GFP positive) were enriched by cell sorting with a BD FACSAria cell sorter. Double TCR gene-transfectants were then created by a second round of transduction of SKW3.LC13β with individual complementary pMIG constructs. Each transfectant line was cloned or enriched using a BD FACSAria cell sorter with staining of anti-CD3ε. The double TCR gene-transfectants include: SKW3.LC13αβ (LC13), SKW3.pre-Tα–LC13β (WT), SKW3.pre-Tα.W46R–LC13β (W46R), SKW3.pre-Tα–LC13β.Y35A (βY35A), SKW3.pre-Tα–LC13β.F108A (βF108A).

Control: the conventional LC13 TCR is specific to FLR epitope (FLRGRAYGL) of Epstein–Barr virus in the context of HLA-B8[8].

Immunofluorescence. Human SKW3 T cells stably expressing WT TCR (TCRα–TCRβ), pre-TCR (pre-Tα–TCRβ) or mutant pre-TCR (pre-Tα–TCRβ-Y35A, pre-Tα-W46R–TCRβ, pre-Tα–TCRβ-F108A) were incubated in RPMI medium containing 1% (v/v) heat-inactivated FBS and left unstimulated or incubated with

mouse anti-human CD3ε (5 μg ml^{-1} OKT3; eBioscience) for 15 min on ice, washed and stimulated by crosslinking with goat anti-mouse immunoglobulin G (20 μg ml^{-1}) at 37 °C for 30 min. Cells were fixed with 0.5% (w/v) paraformaldehyde for 30 min, washed with PBS containing 3% (v/v) FBS and resuspended in PBS containing 5% (v/v) FBS plus 0.1% (w/v) saponin plus primary antibody (OKT3 or anti-TCRβ), and incubated at room temperature for 1 h. Cells were then washed and incubated with the same buffer plus secondary antibody (Alexa Fluor 568 goat anti-mouse; Invitrogen-Molecular Probes) for 30 min. Washed cells were then stained for DNA with Hoechst 33258 (Invitrogen-Molecular Probes) for 10 min, collected onto glass slides by centrifugation (500 g, 5 min) and processed for microscopic analysis using a Zeiss Axioskop 2 MOT Plus epi-fluorescence microscope (Zeiss).

21. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
22. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58,** 1948–1954 (2002).
23. Zwart, P. H. *et al.* Automated structure solution with the PHENIX suite. *Methods Mol. Biol.* **426,** 419–435 (2008).
24. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
25. Cole, J. L. *et al.* Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol.* **84,** 143–179 (2008).
26. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78,** 1606–1619 (2000).
27. Rinner, O. *et al.* Identification of cross-linked peptides from large sequence databases. *Nature Methods* **5,** 315–318 (2008).
28. Burger, R., Hansen-Hagge, T. E., Drexler, H. D. & Gramatzki, M. Heterogeneity of T-acute lymphoblastic leukemia (T-ALL) cell lines: suggestion for classification by immunophenotype and T-cell receptor studies. *Leuk. Res.* **23,** 19–27 (1999).
29. Hirano, T. *et al.* In vitro immune response of human peripheral lymphocytes: IV. Specific induction of human suppressor T cells by an antiserum to the T leukemia cell line HSB. *J. Immunol.* **123,** 1133–1140 (1979).
30. Holst, J., Vignali, K. M., Burton, A. R. & Vignali, D. A. A. Rapid analysis of T-cell selection in vivo using T cell-receptor retrogenic mice. *Nature Methods* **3,** 191–197 (2006).

Author Queries

Journal: **Nature**
Paper: **nature09448**
Title: **The structural basis for autonomous dimerization of the pre-T-cell antigen receptor**

Query Reference	Query
1	AUTHOR: When you receive the PDF proofs, please check that the display items are as follows (doi:10.1038/nature09448): Figs 1, 2, 3, 4, 5 (colour). Please check all figures very carefully as they have been re-labelled, re-sized and adjusted to Nature's style. Please check the title and the first paragraph with care, as they may have been re-written to aid accessibility for non-specialist readers.
2	PROOFREADER: Please update/confirm the tentative publication date.

For Nature office use only:

Layout	☐	Figures/Tables/Boxes	☐	References	☐
DOI	☐	Error bars	☐	Supp info (if applicable)	☐
Title	☐	Colour	☐	Acknowledgements	☐
Authors	☐	Text	☐	Author contribs (if applicable)	☐
Addresses	☐	Methods (if applicable)	☐	COI	☐
First para	☐	Received/Accepted	☐	Correspondence	☐
Display items	☐	AOP (if applicable)	☐	Author corrx	☐