Cytokine 116 (2019) 161–168

Contents lists available at ScienceDirect

Cytokine

journal homepage: www.elsevier.com/locate/cytokine

The tertiary structure of γc cytokines dictates receptor sharing

Pedro A. Reche

Laboratory of Immunomedicine, School of Medicine, Department of Immunology & O2, Universidad Complutense de Madrid, Ave Complutense S/N, Madrid 28040, Spain

ARTICLE INFO

Keywords:
γc cytokines
IL-9
IL-7
Structural comparison
Sequence comparison

ABSTRACT

The γc family of cytokines comprising interleukin-2 (IL-2), IL-4, IL-7, IL-9, IL-15 and IL-2 is an important group of 4-helix bundle cytokines that signals through receptors incorporating the common gamma chain (γc). These cytokines are involved in lymphocyte biology and their specific functions are contingent on binding to cognate receptor chains. Here, we examined the structural relationships between γc cytokines, aiming to understand the basis for receptor chain usage and sharing. To that end, we obtained tertiary structures of human and mouse γc cytokines plus two other related cytokines, IL-13 and TSLP, which share receptors with IL-4 and IL-7, respectively. Subsequently, we compared the cytokine 3D-structures introducing a structural similarity score that grouped γc cytokines in a manner that mirrored the relationships dictated by receptor sharing. Unlike previously thought, we identified that IL-9 is more closely related to IL-2 and IL-15 than to IL-7, which is actually the most distant member of the γc family of cytokines. Moreover, we found that all the members of the γc family of cytokines share the topology of short-chain 4-helix bundle cytokines but IL-7 that with TSLP has the topology of long-chain 4-helix bundle cytokines. We also carried out Maximun-Likehood and Bayesian phylogenetic analyses that supported these results at the amino acid sequence level. Overall, our findings are of paramount relevance to understand receptor sharing among γc cytokines and can lead to the discovery of new cytokine receptor partners.

1. Introduction

Cytokines are a diverse group of cell-signaling soluble factors that play a key role in immunity, regulating the development and activity of many cell types. Cytokines can be classified by different criteria, including function, range of action, cell type that produce them and structure [1]. An important and numerous group of cytokines have a fold consisting of a helical bundle made of four α-helices, designated A, B, C, D, with a unique up-up-down-down topology [2,3]. Four-helix bundle cytokines has been traditionally divided into short-chain and long-chain based on the length of the α-helices and some other structural/topological considerations. In particular, short-chain 4-helix bundle cytokines has an A–B crossover connection that passes behind helix D while the same crossover passes in front of helix D in long-chain cytokines [2,3].

Four-helix bundle cytokines signals through receptor complexes made of one or more distinct class I cytokine receptor chains, also known as hematopoetin receptors. These receptors feature a cytokine-binding homology region (CHR) comprising two Ig-like fibronectin type III (FNIII) domains [4] and are often shared by different cytokines [5]. The best studied paradigm is that of the common γ chain receptor, γc, that it is a shared by interleukin-2 (IL-2), IL-4, IL-7, IL-9, IL-15 and IL-21 [6]. The γc receptor was first described as the IL-2 receptor γ chain (IL-2Rγ) [7] and defects were shown to cause x-linked severe combined immunodeficiency disease (SCID) [8] highlighting its pivotal role. Herein, we will focus on the γc family of cytokines, also known as the IL-2 family.

The cytokines of IL-2 family are essential for the development and survival of lymphocytes [9]. They have been classified within the group of short-chain 4-helix bundle cytokines [10–12] and bind to receptor complexes that include private and shared receptor chains [5,13]. IL-4, IL-7, IL-9 and IL-21 bind to heterodimeric receptors consisting of the γc and IL-4Rα, IL-7Rα, IL-9Rα and IL-21Rα chains, respectively (Fig. 2). In contrast, IL-2 and IL-15 bind to heterotrimers receptors consisting of shared γc and IL-2Rβ chains plus the specific IL-2Rα or IL-15Rα chains (Fig. 2), respectively, which actually do not belong to the family of hematopoietin receptors; they both engage cytokines through sushi domains instead of FNIII domains [14,15]. Receptor sharing is not limited to the γc and IL-2Rβ chains. Interestingly, IL-4 can also bind to a receptor heterodimer consisting of IL-4Rα and IL-13Rα1, which is, in addition, the functional receptor of IL-13 [16,17]. Likewise, IL-7Rα is also the receptor of thymic stromal-derived lymphopoietin (TSLP) in combination with the TSLP Receptor (TSLPR) [18,19] (Fig. 1). Both TSLPR and IL-13Rα1 are more related to γc than to other class I cytokine receptors [11].

Hematopoetin receptors signal through protein tyrosine kinases of

E-mail address: parecheg@med.ucm.es.

https://doi.org/10.1016/j.cyto.2019.01.007
Received 6 November 2018; Received in revised form 3 January 2019; Accepted 7 January 2019
Available online 02 February 2019

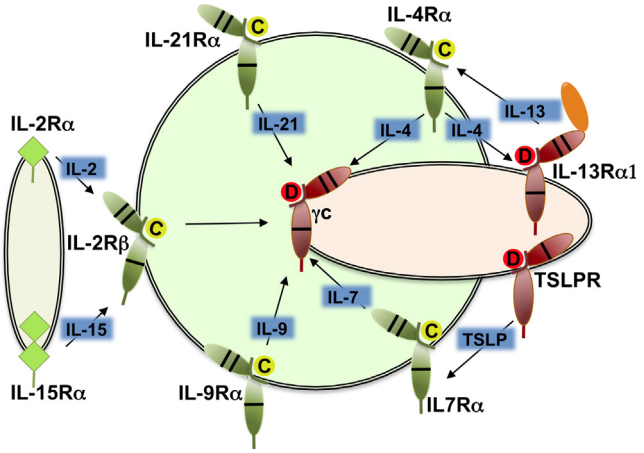

Fig. 1. γc cytokine-receptor interactions. The figure shows γc cytokines and their interactions with their cognate receptors. IL-13 and TSLP share receptors with γc cytokines and are also included. Cytokine-receptors that are connected by a double line and shown in the same color are more closely related to each other than to other receptors and engage the same cytokine helices. γc cytokines bind to their cognates receptors in an elbow shaped region formed by the two conjoined FNIII3 domains in an orderly manner indicated by arrow directions. γc cytokines engage the alpha receptor chains by helix C (yellow), IL2Rβ in the case of IL-2 and IL-15, and then the γc chain by helix D (red). Helix A of the cytokine binds to both receptor chains (not indicated). IL-2 and IL-15 signal through heterotrimeric receptors and in addition to IL2Rβ and γc chains bind to private alpha receptors, IL-2Rα and IL-15Rα, respectively, that do not participate in signaling and are not type I cytokine receptors. As shown in the figure, IL-4 also signals through a heterodimeric receptor in which IL-13Rα1 substitutes the γc chain. This same receptor combination, IL-4Rα and IL-13Rα1, is the functional receptor of IL-13. However, while IL-4 binds first to IL-4Rα and then to γc or IL-13Rα1, IL-13 engage first with IL-13Rα1 and then IL-4Rα. Likewise, TSLP engages first TSLPR and then IL-7Rα. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the Janus family (JAK) that, associated to their cytoplasmic tails, phosphorylate specific STAT (Signal Transducer and Activator of Transcription) proteins. JAK activation and signal transduction occurs when individual receptor chains are brought together thanks to the binding of the cytokines, which act as signaling switches. Biochemical and structural studies have revealed that γc cytokines engage their receptors in an orderly and conserved manner [5,20]. IL-4, IL-7, IL-9 and IL-21 bind first to the α receptors via helix C and A and subsequently bind to the γc chain using mostly helix D [20]. IL-2 and IL-15 binds to IL-2Rβ and γc following this same paradigm but in addition can bind first to the private α receptors by a region comprising helix B [21,22]. Likewise, IL-4 and IL-13 also engage IL-4Rα via helices A and C and IL-13Rα1 with helix D. However, while IL-4 binds first to IL-4Rα and then to IL-13Rα1, IL-13 follows the opposite order [17] Similarly, TSLP bind first to TSLPR receptor through helix D and then to IL-7Rα through helices A and C [23 24].

Despite γ_c cytokines share and engage receptors in a similar manner sequence identity between the members of the family is surprisingly low, under 19%. Such a low sequence identity hinders the identification of functional and evolutionary relationships between these cytokines. Moreover, it supports the presence of additional cytokine receptors. Given the lack of sequence identity, shape complementarity has been to postulated to play a dominant role determining receptor sharing [5]. In this study, we compared the tridimensional (3D)-structures of mouse and human γc families plus mouse and human TSLP and IL-13 (outgroup cytokines) through unsupervised clustering. To this end, we first obtained 3D-structures of all cytokines, completing those with missing residues and modeling the 3D-strucutre of cytokines like IL-9 without solved structures. In addition, we also carried out phylogenetic analyses

Table 1
Uniprot and PDB accession codes of cytokines analyzed in this study.

Cytokines	UNIPROTKB ENTRY	UNIPROTKB AC	PDB	Missing residues[1]
hIL-2	IL2_HUMAN	P60568	1M47	YES
mIL-2	IL2_MOUSE	P04351	4YQX	NO
hIL-9	IL9_MOUSE	P15247	NA	–
mIL-9	IL9_HUMAN	P15248	NA	–
hIL-21	IL21_HUMAN	Q9HBE4	2OQP	
mIL-21	IL21_MOUSE	Q9ES17	NA	–
hIL-7	IL7_HUMAN	P13232	3DI2	YES
mIL-7	IL7_MOUSE	P10168	NA	–
hIL-15	IL15_HUMAN	P40933	2Z3Q	YES
mIL-15	IL15_MOUSE	P48346	2PSM	NO
hIL-4	IL4_HUMAN	P05112	2INT	NO
mIL-4	IL4_MOUSE	P07750	NA	–
hIL-13	IL13_HUMAN	P35225	3LB6	NO
mIL-13	IL13_MOUSE	P20109	NA	–
hTSLP	TSLP_HUMAN	Q969D9	4NN5	YES
mTSLP	TSLP_MOUSE	Q9JIE6	5J11	YES

[1] Whether the selected 3D-structure have missing intercalating residues. NA: Not available.

upon a structure-guide multiple sequence alignment of the mentioned cytokines. In contrast to the current view, structural and sequence comparisons indicated that IL-9 is closer to IL-2 and IL-15 than to IL-7. Moreover, we found that all γc cytokines share the topology of short-chain 4-helix bundle cytokines but IL-7, which resemble that of long-chain cytokines. These results emphasize the value of structural studies in determining relationships between cytokines and can lead to unravel new functional cytokine receptor complexes.

2. Material and methods

2.1. Cytokine sequences and 3D-structures

Mouse and human IL-2, IL-4, IL-7, IL-9, IL-15, IL-13 and TSLP were obtained from the UNIPROTKB database [25]. We identified available representative 3D-structures of these cytokines upon BLAST searches [26] against the Brookhaven Protein Data Bank (PDB) subset of protein sequences at the NCBI. UNIPROTKB accession numbers and representative PDB codes of the cytokines used in this study are indicated in Table 1 (we did not considered PDBs corresponding to mutant cytokine sequences). Subsequently, we retrieved the relevant PDB files and used RASMOL [27] to visualize all structures and verify those with missing residues in connecting loops (indicated in Table 1).

2.2. Tertiary-structure modeling

We used the I-TASSER server [28] to model the 3D-structure of cytokines without solved 3D-strucure, including mouse IL-9 and human IL-9, and to complete the 3D-structure of cytokines with missing residues in connecting loops. In addition, we also completed the 3D-structure of mouse IL-2 (PDB: 4YQX) adding 7 Gln residues at the N-terminus thus making helix A comparable in length to that of human IL-2. I-TASSER uses iterative threading reassemble refinement and consistently ranks among the top 3D-strucuture modeling methods in Critical Assessment of protein Structure Prediction (CASP) competitions [29] For those cytokines without any 3D-structure available, we entered in the server the sequence corresponding to the matured cytokine polypeptide as identified in UNIPROTKB. For those cytokines with incomplete 3D-strucutres, we entered in the server the amino acid sequence obtained after the relevant PDBs completed with the missing residues. We only considered the top ranking model reported by I-TASSER according to computed *c*-scores. The *c*-score is a confidence measure of the quality of predicted models by I-TASSER. The values of *c*-score typically range between −5 and 2 with high values translating in high confidence models.

Table 2
Analysis of cytokine 3D-structures used in this study.

Cytokines	PDB[1]	I-TASSER[2] *c*-scores	Modeling Top Template	Residues	Ramachandran Plot Statistics % FR	% DR
hIL-2	1M47(X-RAY)	1.49	1irlA (hIL-2)	S6–L132	98.4	1.7
mIL-2	4YQX (X-RAY)	1.03	4yqxA (mIL-2)	Q20–S147	96.0	0.8
hIL-9	NA	−1.39	2z3qA (hIL-15)	G2-G121	92.6	3.7
mIL-9	NA	−1.24	2z3rA (hIL-15)	R2-R121	91.7	4.6
hIL-21	2OQP (NMR)	–	–	R6–L124	95.4	1.9
mIL-21	NA	1.25	2oqpA (hIL-13)	R10–L128	96.2	1.9
hIL-7	3DI3 (X-RAY)*	0.02	3di3A (hIL-7)	K7-M147	97.7	0.8
mIL-7	NA	0.5	3di3A (hIL-7)	K7-K126	95.6	1.8
hIL-15	2Z3Q (X-RAY)*	1.52	2z3rO (hIL-15)	N1-S114	99.1	0
mIL-15	2PSM (X-RAY)	–	–	N1-S114	100	0
hIL-4	2INT (X-RAY)	–	–	C3-C127	98.3	0.8
mIL-4	NA	1.28	1ITI (hIL-4)	C5-S120	98.2	0.9
hIL-13	3LB6 (X-RAY)*	–	–	T8-G110	100	0
mIL-13	NA	0.81	1ga3A (hIL-13)	L10-G108	100	0
hTSLP	5J11*	0.93	5j11A (hTSLP)	D7-R125	96.3	3.6
mTSLP	4NN5*	1.18	4NN5 (mTSLP)	N7-Q118	99.0	0

[1] PDB codes with representative cytokine 3D-structures. NA, Not Available.
* PDB with missing residues in connecting regions.
[2] Cytokines without solved structure or missing some residues in the solved structures were subjected to modeling using I-TASSER. We also model the 3D-structure of mIL-2 to include 6N-terminal Qln residues. We show *c*-scores of top models (model 1) and PDB codes with chain id (capital letter) of top-templates used by I-TASSER (in bracket the cytokine corresponding to the PDB). Solved 3D-structure and theoretical models were evaluated using Ramachandran plots and we indicate the fraction of residues that showed in favored regions (%FR) and in disallowed regions (%DR). In column "*Residues*" we show the cytokine region that was considered for PROCHECK analysis and further structural comparisons. Cytokine PDBs used for structural comparisons are provided in Supporting Material 1.

2.3. Comparison of cytokine 3D-structures

Prior to comparing the cytokine 3D-structures, we inspected all the 3D-structures using RASMOL [27], both modeled or experimental, and discard N-terminal and C-terminal residue extensions spanning beyond the first and last helix of the 4-helical bundle. Subsequently, we carried pairwise superimpositions/alignments of all-*vs*-all cytokine 3D-structures using DaliLite [30] and built two different types of square matrices after Z_{ij} and S_{ij} coefficients, respectively. Z_{ij} matrix coefficients consisted of Z-scores reported by DaliLite resulting from aligning structure i and j, and S_{ij} represented the percentage of aligned residues (S-scores) as computed by Eq. (1).

$$S_{ij} = 100 * \frac{A_{ij}}{L_{ij}} \quad (1)$$

were *i, j*, represent cytokines i and j, whose 3D-strutures has been aligned, A_{ij} is the number of aligned residues reported by DaliLite and L_{ij} is s the shortest of the sequence length of cytokines i and j. Subsequently, we used the R-package PVCLUST [31] to cluster the cytokine 3D-structures upon the described matrices. PVCLUST reports for each cluster two probability values, an approximately unbiased (AU) *p*-value and a bootstrap probability (BP) value, that are computed upon bootstrap analyses. Of those two, we chose to show *p*-values to assess branch confidence. We selected euclidean distances and ward method [32] to generate clusters. Cladograms generated by PVCLUST were converted to phylograms using the R-package APE [33]. The author will provide upon written request the R code used to generate cladrograms and phylograms.

2.4. Multiple sequence alignments and phylogenetic analysis

We obtained a structure-guided multiple sequence alignment (MSA) of mouse and human IL-2 cytokine family members and IL-13 and TLSP using PROMALS3D [34]. We executed PROMALS3D with the Dali option which utilize DaliLite [30] for 3D-structure guidance. For these alignments, we only aligned the portion of the cytokine sequences containing the 4-helix bundle identified from the relevant 3D-structures as indicated elsewhere.

To analyze cytokine sequence kinship, we produced phylogenetic trees upon the generated MSA using CLUSTLAX [35], IQTREE [36] and MrBayes [37]. In CLUSTLAX we produced bootstrap neighbor-joining [38] trees using 1000 replications to address branch confidence [39]. IQTREE estimates phylogeny using Maximun-Likehood [40] and we used the software to generate optimal tress combining the model finder option with bootstrapping (*'-m TEST -alrt 1000 -bb 1000'*). We used MrBayes to perform bayesian phylogenetic inference [41] using the following parameters. We set the priors to a fixed "wag" amino acid rate matrix model (*prset aamodelpr = fixed(wag)*), the likelihood model to a gamma distribution with invariable sites and 4 categories (*lset rates = invgamma Ngammacat = 4*), and carried Markov chain Monte Carlo (MCMC) simulations over five runs of 500,000 generations, with sampling frequency of 10 (*mcmc nchains = 4 ngen = 500000 nruns = 5 printfreq = 1000 samplefreq = 10 savebrlens = yes starttree = random*). Branch support was evaluated by posterior probabilities after a burn-in of 25%. After the simulations the topological convergence diagnostic was lower than 0.005.

2.5. Other procedures

We used PyMOL Molecular Graphics System for generating molecular renderings and DSSP [42] for secondary structure assignments. To evaluate quality of 3D-structures, we used PROCHECK [43] at the SAVES site: http://servicesn.mbi.ucla.edu/SAVES/. We used FigTree for drawing phylogenetic trees (http://tree.bio.ed.ac.uk/software/figtree/).

3. Results and discussion

3.1. Generation of 3D-structure models of γc cytokines and related cytokines sharing receptors chains

To investigate structural relationships among γc cytokines we also considered TSLP and IL-13, as they share receptors with IL-7 and IL-4. All these cytokines have been extensively studied and we sought for human and mouse solved 3D-structures. We identified representative 3D-structures for all the cytokines but IL-9, whose structure remains to be determined. IL-21, IL-7 and IL-4 3D-structures were available for human but not mouse. In addition, some of the identified cytokine

structures, including those of human IL-15 and IL-7, human and mouse TSLP and mouse IL-2 have missing residues in connecting regions or were incomplete. Subsequently, we used I-TASSER to generate 3D-structure models for all the relevant cytokines with either no solved 3D-structures or solved 3D-structures with missing residues (details in Material and Methods).

As expected I-TASSER built cytokine 3D-models after counterparts with representative 3D-structure when available (Table 2). Thus, I-TASSER generated 3D-structures of mouse IL-21, IL-7 and IL-4 upon the know 3D-structures of human IL-21, IL-7 and IL-4 respectively. Likewise, I-TASER completed the 3D-structures of cytokines with missing residues upon the corresponding solved 3D-structures. Models' confidence values (*c*-scores) were quite high in these two scenarios and ranged from 0.02 (hIL-7) to 1.52 (mIL-13). Interestingly, human and mouse IL-7 models had the lowest *c*-scores, 0.02 and 0.5, respectively, of all the models generated for cytokines with available representative 3D-structures. Note that we generated 3D-models for mouse and human IL-7 since the available 3D-structures missed some residues. Overall, however, the lowest *c*-scores corresponded to the models generated for human and mouse IL-9 (Table 2).

The human and mouse IL-9 models generated by I-TASER had confidence values of −1.39 and −1.24, respectively. Interestingly, the top template 3D-structures picked by I-TASER to model the 3D-structure of IL-9 corresponded to IL-15 (Table 2) while IL-7 was not considered. This result was both interesting and unexpected as IL-9 and IL-7 had long been considered within the same family [44] (INTERPRO family IPR000226). Given that *c*-scores of the IL-9 models were somewhat low, yet above the median value of the *c*-score range (−5, 2), we evaluated the quality of the IL-9 models using PROCHECK and found that it was comparable with both, that of the remaining models and that of the experimental 3D-structures (Table 2). Thus, we believe that 3D-strucutre of IL-9 modeled here is reasonably correct. The 3D-coordinates of hIL-9 and mIL-9 are provided in Supporting Material 1.

The predicted 3D-structure of IL-9 is that of a short-chain 4-helical bundle cytokines, in which the segment connecting helices A and B pass behind helix D (Fig. 2). Some short-chain 4-helical bundle cytokines often bear a two-stranded β-sheet where the two connecting crossovers meet [2] that is absent in IL-9 models (Fig. 2). However, a secondary structure analysis of the IL-9 models using DSSP finds β-strand propensity in that region (data not shown) and it cannot be discarded that a two-stranded β-sheet will turn out in that region once the structure is solved.

3.2. Structural comparison of γc cytokines and related cytokines sharing receptors chains

To analyze kinship between γc cytokines, we carried out all-*vs*-all pairwise cytokine 3D-structure alignments and obtained structural similarity matrices after Z-scores and S-scores (details in Material and Methods). Z- and S-score matrices are provided in Supporting Material 2 and 3, respectively. Subsequently, we carried an unsupervised clustering analysis with statistics support using PVCLUST as detailed in Material and Methods. The results of such analysis are shown in Fig. 3.

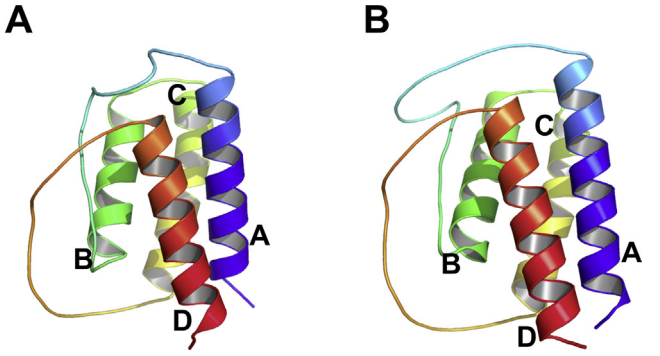

Fig. 2. Tertiary structure of IL-9. Ribbon renderings of the 3D-structures obtained for human (A) and mouse (B) IL-9. Note the typical fold of short-chain 4-helical cytokines in which the segment connecting helices A and B pass behind helix D. The estimated root mean square deviations (RMSD) of the hIL-9 and mIL-9 3D-models were 7.3 ± 4.2 Å and 7.0 ± 4.1 Å, respectively.

The phylogram obtained after S-cores (S-phylogram, Fig. 3A) divides γc cytokines in two groups, one including IL-7 that clusters/branches with TSLP and another including the remaining γc cytokines and IL-13 (*p*-value = 100). In this second cluster, IL-21 splits apart from IL-2, IL-4, IL-9 and IL-15 that cluster together (*p*-value = 82). Within this group, IL-4 clusters with IL-13 with great confidence (*p*-value = 93) and so IL-9 with IL-2 and IL-19 (*p*-value = 95). This clustering schema is somewhat different from that obtained after Z-scores (Z-phylogram, Fig. 3B). In the Z-phylogram, there are also two main branches. The branch with greater confidence value (*p*-value = 89) includes IL-21 along with IL-2, IL-9 and IL15 that again branch together (*p*-value = 88). However, IL-9 is closer to IL-15 than to IL-2 in this phylogram, while in the previously described S-phylogram IL-9 is closer to IL-2. However, the major different between the Z-phylogram and the S-phylogram lies in the second main branch, which, in the Z-phylogram, includes IL-4 and IL-13 along with IL-7 and TSLP. Moreover, IL-7 appears closer to IL-4 and IL-13 (*p*-value = 81) than to TSLP that appears as the most distant cytokine in this group.

In sum, both, Z-scores and S-scores led to the clustering of IL-9 with IL-2 and IL-15. However, the S-phylogram obtained upon clustering S-scores (percentage of 3D-aligned residues) better reflects the know relationships between γc cytokines and other related cytokines that are dictated by receptor sharing. This was confirmed upon inspecting the topology of all γc cytokines, as we found that IL-7 is indeed different from the remaining cytokines. Thus, all γc cytokines but IL-7 share the topology of short-chain 4-helex bundle cytokines (Fig. 4A). Surprisingly, IL-7 and TSLP exhibit the typical topology of long-chain 4-helex bundle cytokines (Fig. 4B), in which the crossover connecting helix A and B pass in front of helix D. This unusual topology of IL-7 and TSLP has not been reported earlier, probably due the fact that connecting loops are missing from the reported 3D-structures [23,24,45].

It is worth nothing that there are some R-packages such as Bio3D [46] and online servers such as DALI [47] that can be used to infer phylogenetic relationships from 3D-structures. However, these tools do not provide confidence values for branches and, in our hands, did not reproduce the known relationships existing between γc cytokines (data not shown).

3.3. Amino acid sequence analysis of γc cytokines

We carried sequence comparison analysis of human and human γc cytokines to check if the unexpected kinships observed at the structural level could also be verified at the protein sequence level. To that end, we extracted the amino acid sequences from the cytokine 3D-structures (Table 2) and produced a structure-guided multiple sequence alignment (MSA) of γc cytokines, TSLP and IL-13 (details in Material and Methods). The resulting MSA (Fig. 5A) does not show a single conserved residue across all the cytokines and exhibits multiple gaps; yet, it clearly reveals 4-ungapped blocks corresponding to the conserved α-helices. After the MSA, we computed the sequence identity between all cytokines (Fig. 5B) and generated a neighbor-joining phylogenetic tree with 1,000 bootstrapping replications using CLUSTALX (Fig. 5C).

Sequence identity among γc cytokines is very low and range from 56 ± 9% for orthologs to 10.3 ± 1.6% between different human γc cytokines. The largest and lowest identity between human and mouse cytokine orthologs is for IL-15 (70.17%) and TLSP (36.6%), respectively. On the other hand, the largest and lowest identity among distinct human cytokines considered in this study is found between IL-15 and IL-2 (19.3%) and between TSLP and IL-13 (8.73%), respectively (Fig. 5B). The CLUSTALX phylogram (Fig. 5C) correctly identifies the

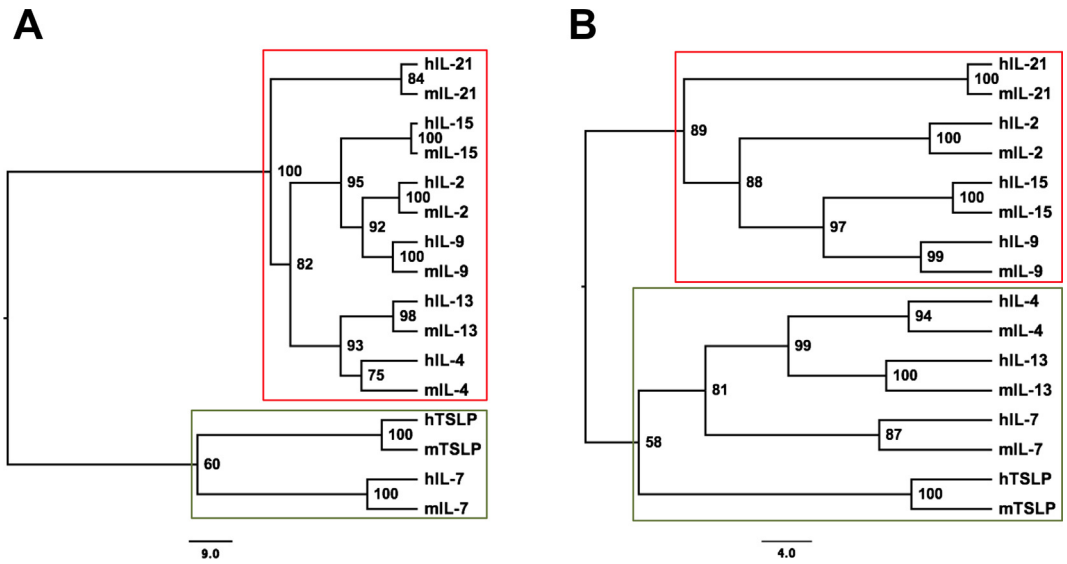

Fig. 3. Structural kinship among γc cytokines. Phylograms generated upon Z-scores (A) and percentage of aligned residues, S-scores (B) obtained from all-*vs*-all pairwise alignments of cytokine 3D-structures. Node confident values consist of unbiased *p*-values computed after bootstrap analyses with PVCLUST (details in Material and Methods). Green and red squares highlight clusters including TSLP and IL-9, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

relationship between IL-2 and IL-15 and that between IL-4 and IL-13 (bootstrap values ≥500). Moreover, it could be seen to support that IL-9 is more closely related to IL-7 (braches with TSLP) than to other γc cytokines. However, the confidence of this relatedness is low (bootstrap value 376) and it is the same as that between IL-9 and the remaining γc cytokines. Neighbor-joining phylogenetic trees generated with CLUSTALX represents a quick and valuable way to assess phylogeny. However, given that branching of IL-9 with IL-7 was marginally supported and contradicted our previous results, we subjected the MSA to Maximum-Likelihood (ML) and Bayesian phylogenetic analysis (details in Material and Methods).

The phylograms obtained by ML and Bayesian inference (Fig. 6) are quite similar and, in contrast to the CLUSTALX phylogram, fully support the kinship of γc cytokines unraveled at the structural level (Fig. 3). Thus, IL-7 is clearly closer to TSLP than to other γc cytokines: *p*-value of 75 in the ML phylogram (Fig. 6A) and of 87 in the Bayesian phylogram (Fig. 6B). In addition, IL-9 branches with IL-2, IL-15 and IL-21 with reasonable confidence in the ML phylogram (*p*-value of 58) and good confidence in the Bayesian phylogram (*p*-value 75). In contrast to structural-based phylograms, sequence-based phylograms appear to support that IL-21, rather than IL-9, to be more closely related to IL-15 and IL-2. However, such grouping seems unreliable (*p*-value ≤45) in comparison with that between IL-9, IL-2 and IL-15 (*p*-value ≥85) observed in the structural phylograms (Fig. 3).

3.4. Functional implications of the unraveled kinship among γc cytokines

There are two major findings of this study that can both lead to discover new cytokine-receptor partnerships. One is the distant structural relationship of IL-7 with the remaining γc cytokines and its close

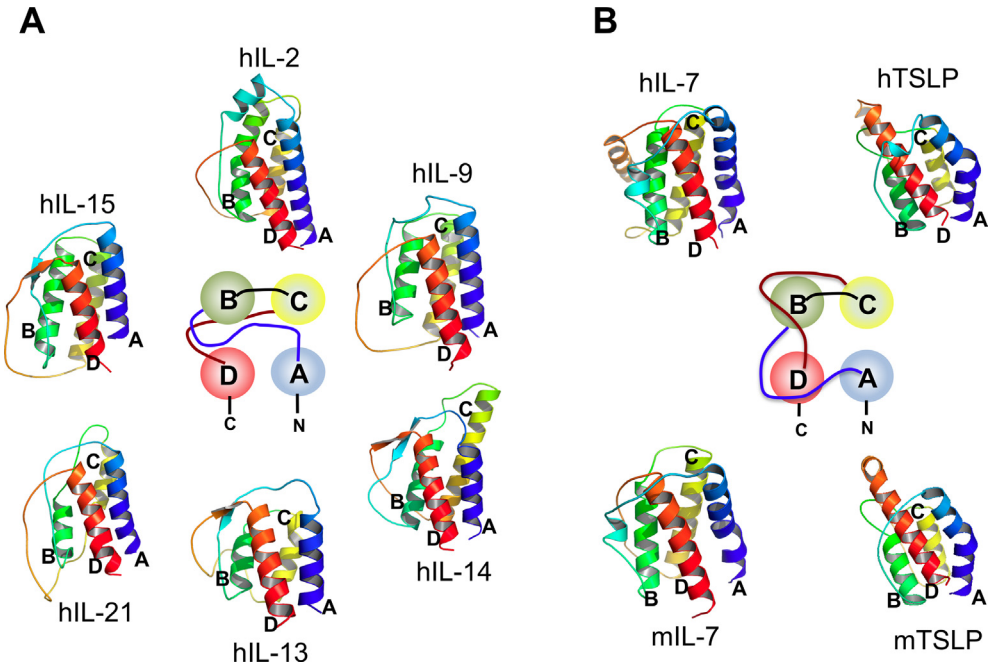

Fig. 4. Structural diversity in γc cytokines. (A) Ribbon rendering of the 3D-structure of human IL-2, IL-9, IL-15, IL-21, IL-4 and IL-13. These cytokines have the typical fold topology of short-chain 4-helical bundle cytokines (center of the panel), in which the segment connecting helix A (blue) and B (green) is behind both helix D (red) and the segment connecting helix C (yellow) and D, as shown in the schematic drawing of the topology. (B) Ribbon rendering of human and mouse IL-7 and TSLP. These cytokines share the fold topology of long-chain 4-helical bundle cytokines, in which the segment connecting helix A (blue) and B (green) pass in front of helix D. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

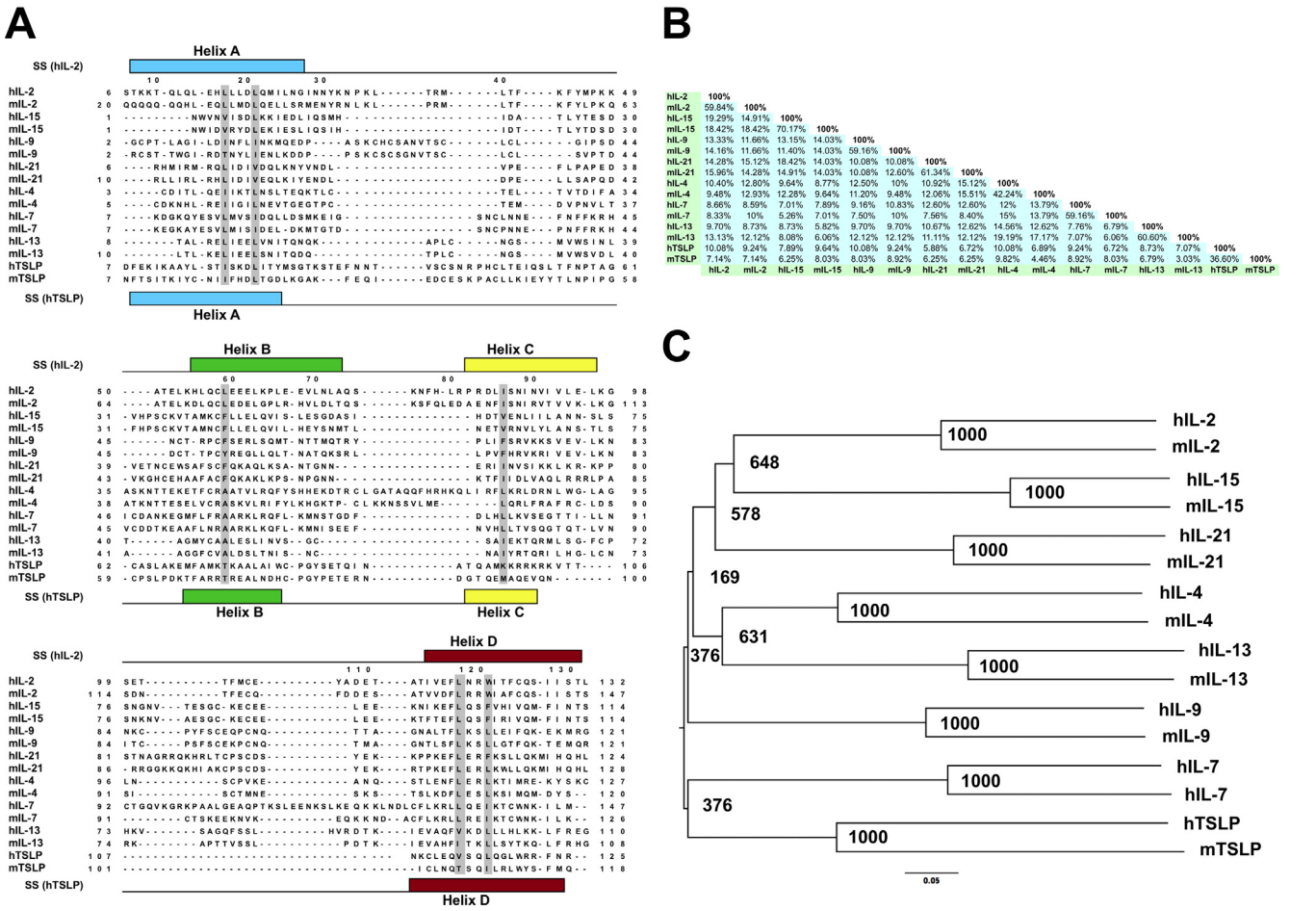

Fig. 5. Sequence amino acid comparison of γc cytokines. (A) Structure-guided MSA of human and mouse γc cytokines plus human and mouse IL-13 and TSLP. Numbering on top of the MSA corresponds to that of the mature hIL-2. The conserved secondary structure (SS) helices A (blue), B (green), C (yellow) and D (red) depicted on top and under the MSA correspond to hIL-2 (PDB: 1M47) and hTLSP (PDB: 4NN5). There is not a single amino acid residue strictly conserved among these cytokines but each helix contain one or two positions consisting mostly of hydrophobic residues, shadowed in grey. MSA is provided in supporting Material 4 (B) Diagonal of symmetric matrix depicting the percentage of identity between cytokines. Identities were computed from the MSA in panel A using the SIAS server at http://imed.med.ucm.es/Tools/sias.html. (C) Neighbor-joining tree built upon the structure-based MSA shown in panel A using CLUSTALX. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

proximity to TSLP. In contrast to all γc cytokines and IL-13, IL-7 and TSLP have both a fold topology that resembles that of long-chain chain 4-helix bundle cytokines (Fig. 4). To our knowledge, this unique 3D-structure topology of IL-7 and TSLP has not been noticed before, surely due to the fact that crossover connections are missing from the available 3D-structures. In any case, this finding highlights the ability of hematopoietin receptors, as represented by the γc chain, to engage structurally diverse cytokines, and it supports the existence of additional cytokine receptor chain sharing, yet to be uncovered, between short-chain and long-chain 4-helix bundle cytokines.

The second major finding of this study is the divorce of IL-9 with IL-7. Structurally, IL-9 is much related to other γc cytokines, in particular IL-2 and IL-15, than to IL-7 (Fig. 4). This structural relationship is somewhat expected as I-TASSER built the 3D-structure of IL-9 after 10 templates and 7 corresponded to IL-15. The remaining templates corresponded to IL-4 (PDB: 1ITI), IL-21 (PDB: 2OQP) and FLT3 ligand (PDB: 1ETE). Nonetheless, since receptor usage and structure of γc cytokines are connected, it is tantalizing to speculate that, to hold true, the close structural relationship between IL-9, IL-2 and IL-15 may anticipate new receptor partners. Thus, we wonder if IL-9, similarly to IL-2 and IL-15 [5], could also engage an additional receptor, related to IL2-Rα and IL-15Rα (Fig. 2), to form a heterotrimeric receptor with the IL-9Rα and γc chains. In fact, such additional receptor chain could perhaps be the IL-2Rα or the IL-15Rα chains. In this regard, the STRING database [50] shows a functional link between IL-15Rα and IL-9 (ID: 9606.ENSP00000244174), but not between IL-2Rα and IL-9 (ID: 9606.ENSP00000369293). Furthermore, we also wonder if IL-9 could signal through receptors complexes encompassing IL-2Rβ in addition to the γc chain and likewise if IL-2 and/or IL-15 could signal through receptors complexes including the IL-9Rα instead of the IL-2Rβ. In this particular, the STRING database identifies functional ties between both, IL-2 and IL-15, and IL-9R (ID: 9606.ENSP00000244174). Functional relationships found by STRING do not necessarily translate into protein-protein interactions and clearly receptor sharing between IL-9 and IL-2 and IL-15, if any, will need to be proved.

4. Conclusions

Receptor chain usage and sharing among γc cytokines comprising IL-2, IL-4, IL-7, IL-9, IL-15 and IL-21 is key to understand their biological function and it is thought to be determined by shape complementarity [5]. Here, we anticipated receptor sharing among γc cytokines through 3D-structure comparisons. Unlike previously reported, we identified that IL-9 is more closely related to IL-2 and IL-15 than to IL-7, which could also be confirmed through sequence-based phylogenetic analyses. In addition, we also identified IL-7 as the most distant member of the γc family of cytokines and discovered that it has a topology, shared with TSLP, that resembles that of long-chain 4-helix

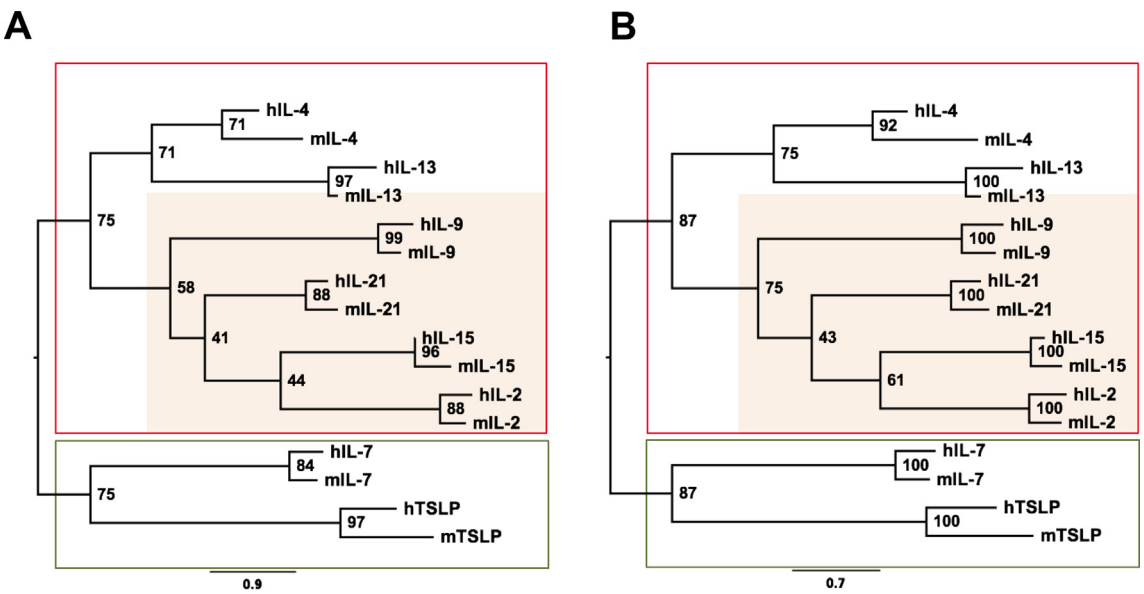

Fig. 6. Amino acid sequence kinship among γc cytokines. Maximun-Likehood (A) and Bayesian (B) phylogenetic trees generated upon the MSA of human and mouse γc cytokines plus IL-13 and TSLP shown in Fig. 6. The Maximun-Likehood phylogram was generated using IQTREE (See Material and Methods) after the best-fitting model identified by the program consisting of a revised general matrix [48] with empirical amino acid frequencies from data and a rate heterogeneity across sites determined by a discrete gamma model with default 4 rate categories (JTTDCMut + F + G4). The Bayesian phylogram was generated using MrBayes with a selected "wag" amino acid matrix [49] and rate heterogeneity across sites determined by an inverse gamma model with 4 rate categories.

bundle cytokines. The close structural relationship between IL-9 and IL-15 and IL-2 suggests that they could exchange receptors, while binding of IL-7 to the γc chain raises the possibility of further receptor chain sharing between short-chain and long-chain 4-helix bundle cytokines.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Pedro A. Reche: Conceptualization, Formal analysis, Investigation, Methodology, Software, Writing - original draft, Writing - review & editing.

Acknowledgements

PAR wishes to thank Francisco Martin Martin & Esther Lafuente Duarte for critical and positive feedback. PAR also thanks the Spanish Department of Science for continuous support through research grants (SAF2006:07879, SAF2009:08301 & BIO2014:54164-R).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cyto.2019.01.007.

References

[1] A. Thomson, M. Lotzke, The Cytokine Handbook, Academic, London/San Diego, 2003.
[2] D.A. Rozwarski, A.M. Gronenborn, G.M. Clore, J.F. Bazan, A. Bohm, A. Wlodawer, et al., Structural comparisons among the short-chain helical cytokines, Structure 2 (1994) 159–173.
[3] N.A. Nicola, D.J. Hilton, General classes and functions of four-helix bundle cytokines, Adv. Protein Chem. 52 (1998) 1–65.
[4] J.F. Bazan, Structural design and molecular evolution of a cytokine receptor superfamily, Proc. Natl. Acad. Sci. USA 87 (1990) 6934–6938.
[5] X. Wang, P. Lupardus, S.L. Laporte, K.C. Garcia, Structural biology of shared cytokine receptors, Annu. Rev. Immunol. 27 (2009) 29–60, https://doi.org/10.1146/annurev.immunol.24.021605.90616 60.
[6] W. Liao, J.X. Lin, W.J. Leonard, IL-2 family cytokines: new insights into the complex roles of IL-2 as a broad regulator of T helper cell differentiation, Curr. Opin. Immunol. 23 (2011) 598–604, https://doi.org/10.1016/j.coi.2011.08.003 Epub Aug 31.
[7] T. Takeshita, H. Asao, K. Ohtani, N. Ishii, S. Kumaki, N. Tanaka, et al., Cloning of the gamma chain of the human IL-2 receptor, Science 257 (1992) 379–382.
[8] M. Noguchi, H. Yi, H.M. Rosenblatt, A.H. Filipovich, S. Adelstein, W.S. Modi, et al., Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans, Cell 73 (1993) 147–157.
[9] J.X. Lin, W.J. Leonard, The common cytokine receptor gamma chain family of cytokines, Cold Spring Harb. Perspect. Biol. (2017) 16.
[10] W.J. Leonard, Cytokines and immunodeficiency diseases, Nat. Rev. Immunol. 1 (2001) 200–208, https://doi.org/10.1038/35105066.
[11] J.L. Boulay, J.J. O'Shea, W.E. Paul, Molecular phylogeny within type I cytokines and their cognate receptors, Immunity 19 (2003) 159–163.
[12] J.L. Boulay, W.E. Paul, Hematopoietin sub-family classification based on size, gene organization and sequence homology, Curr. Biol. 3 (1993) 573–581.
[13] R. Spolski, D. Gromer, W.J. Leonard, The gamma c family of cytokines: fine-tuning signals from IL-2 and IL-21 in the regulation of the immune response, 1000Res. 6 (2017) 1872, https://doi.org/10.12688/f1000research.202.1 eCollection 2017.
[14] M. Rickert, X. Wang, M.J. Boulanger, N. Goriatcheva, K.C. Garcia, The structure of interleukin-2 complexed with its alpha receptor, Science 308 (2005) 1477–1480, https://doi.org/10.126/science.1109745.
[15] I. Lorenzen, A.J. Dingley, Y. Jacques, J. Grotzinger, The structure of the interleukin-15 alpha receptor and its implications for ligand binding, J. Biol. Chem. 281 (2006) 6642–6647, https://doi.org/10.1074/jbc.M513118200 Epub 2005 Dec 23.
[16] A.L. Andrews, J.W. Holloway, S.T. Holgate, D.E. Davies, IL-4 receptor alpha is an important modulator of IL-4 and IL-13 receptor binding: implications for the development of therapeutic targets, J. Immunol. 176 (2006) 7456–7461.
[17] S.L. LaPorte, Z.S. Juo, J. Vaclavikova, L.A. Colf, X. Qi, N.M. Heller, et al., Molecular and structural basis of cytokine receptor pleiotropy in the interleukin-4/13 system, Cell 132 (2008) 259–272, https://doi.org/10.1016/j.cell.2007.12.030.
[18] P.A. Reche, V. Soumelis, D.M. Gorman, T. Clifford, M. Liu, M. Travis, et al., Human thymic stromal lymphopoietin preferentially stimulates myeloid cells, J. Immunol. 167 (2001) 336–343.
[19] L.S. Park, U. Martin, K. Garka, B. Gliniak, J.P. Di Santo, W. Muller, et al., Cloning of the murine thymic stromal lymphopoietin (TSLP) receptor: formation of a functional heteromeric complex requires interleukin 7 receptor, J. Exp. Med. 192 (2000) 659–670.
[20] J.B. Spangler, I. Moraga, J.L. Mendoza, K.C. Garcia, Insights into cytokine-receptor interactions from cytokine engineering, Annu. Rev. Immunol. 33 (2015) 139–167, https://doi.org/10.1146/annurev-immunol-032713-120211 Epub 2014 Dec 10.
[21] X. Wang, M. Rickert, K.C. Garcia, Structure of the quaternary complex of interleukin-2 with its alpha, beta, and gammac receptors, Science 310 (2005) 1159–1163, https://doi.org/10.26/science.1117893.
[22] A.M. Ring, J.X. Lin, D. Feng, S. Mitra, M. Rickert, G.R. Bowman, et al., Mechanistic and structural insight into the functional dichotomy between IL-2 and IL-15, Nat.

Immunol. 13 (2012) 1187–1195, https://doi.org/10.038/ni.2449 Epub 012 Oct 28.
[23] K. Verstraete, F. Peelman, H. Braun, J. Lopez, D. Van Rompaey, A. Dansercoer, et al., Structure and antagonism of the receptor complex mediated by human TSLP in allergy and asthma, Nat. Commun. 8 (2017) 14937, https://doi.org/10.1038/ncomms14937.
[24] K. Verstraete, L. van Schie, L. Vyncke, Y. Bloch, J. Tavernier, E. Pauwels, et al., Structural basis of the proinflammatory signaling complex mediated by TSLP, Nat. Struct. Mol. Biol. 21 (2014) 375–382, https://doi.org/10.1038/nsmb.2794 Epub 014 Mar 16.
[25] M. Magrane, UniProt Knowledgebase: a hub of integrated protein data, Database (Oxford) (2011), https://doi.org/10.1093/database/bar009 2011:bar009, Print 2011.
[26] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, et al., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[27] A. Sayle, E.J. Milner-White, RASMOL: biomolecular graphics for all, Trends Biochem. Sci. 20 (1995) 374.
[28] A. Roy, A. Kucukural, Y. Zhang, I-TASSER: a unified platform for automated protein structure and function prediction, Nat. Protoc. 5 (2010) 725–738, https://doi.org/10.1038/nprot.2010.5 Epub Mar 25.
[29] J. Moult, K. Fidelis, A. Kryshtafovych, T. Schwede, A. Tramontano, Critical assessment of methods of protein structure prediction (CASP)-Round XII, Proteins 86 (2018) 7–15, https://doi.org/10.1002/prot.25415 Epub 2017 Dec 15.
[30] L. Holm, S. Kaariainen, P. Rosenstrom, A. Schenkel, Searching protein structure databases with DaliLite vol 3, Bioinformatics 24 (2008) 2780–2781, https://doi.org/10.1093/bioinformatics/btn507 Epub 2008 Sep 25.
[31] R. Suzuki, H. Shimodaira, Pvclust: an R package for assessing the uncertainty in hierarchical clustering, Bioinformatics 22 (2006) 1540–1542, https://doi.org/10.093/bioinformatics/btl117 Epub 2006 Apr 4.
[32] F. Murtagh, P. Legendre, Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? J. Classif. 31 (2014) 274–295.
[33] E. Paradis, J. Claude, K. Strimmer, APE: analyses of phylogenetics and evolution in R language, Bioinformatics 20 (2004) 289–290.
[34] J. Pei, N.V. Grishin, PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information, Methods Mol. Biol. 1079 (2014) 263–271, https://doi.org/10.1007/978-1-62703-646-7_17.
[35] J.D. Thompson, T.J. Gibson, D.G. Higgins, Multiple sequence alignment using ClustalW and ClustalX, Curr. Protoc. Bioinform. (2002), https://doi.org/10.1002/0471250953.bi0203s00 Chapter: Unit 2.3.
[36] L.T. Nguyen, H.A. Schmidt, A. von Haeseler, B.Q. Minh, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies, Mol. Biol. Evol. 32 (2015) 268–274, https://doi.org/10.1093/molbev/msu300 Epub 2014 Nov 3.
[37] F. Ronquist, M. Teslenko, P. van der Mark, D.L. Ayres, A. Darling, S. Hohna, et al., MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space, Syst. Biol. 61 (2012) 539–542, https://doi.org/10.1093/sysbio/sys029 Epub 2012 Feb 22.
[38] S.M. Firestine, A.E. Nixon, S.J. Benkovic, Threading your way to protein function, Chem. Biol. 3 (1996) 779–783.
[39] J. Felsenstein, Confidence limits on phylogenies: an approach using the bootstrap, Evolution 39 (1985) 783–791.
[40] E.E. Schadt, J.S. Sinsheimer, K. Lange, Computational advances in maximum likelihood methods for molecular phylogeny, Genome Res. 8 (1998) 222–233.
[41] B. Mau, M.A. Newton, B. Larget, Bayesian phylogenetic inference via Markov chain Monte Carlo methods, Biometrics 55 (1999) 1–12.
[42] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637, https://doi.org/10.1002/bip.360221211.
[43] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J. Thornton, PROCHECK – a program to check the stereochemical quality of protein structures, J. Appl. Cryst. 8 (1993) 9.
[44] A.L. Morris, M.W. MacArthur, E.G. Hutchinson, J.M. Thornton, Stereochemical quality of protein structure coordinates, Proteins 12 (1992) 345–364, https://doi.org/10.1002/prot.340120407.
[45] C.A. McElroy, J.A. Dohm, S.T. Walsh, Structural and biophysical studies of the human IL-7/IL-7Ralpha complex, Structure 17 (2009) 54–65, https://doi.org/10.1016/j.str.2008.10.019.
[46] L. Skjaerven, X.Q. Yao, G. Scarabelli, B.J. Grant, Integrating protein structural dynamics and evolutionary analysis with Bio3D, BMC Bioinf. 15 (2014) 399, https://doi.org/10.1186/s12859-014-0399-6.
[47] L. Holm, L.M. Laakso, Dali server update, Nucleic Acids Res. 44 (2016) W351–W355, https://doi.org/10.1093/nar/gkw357 Epub 2016 Apr 29.
[48] C. Kosiol, N. Goldman, Different versions of the Dayhoff rate matrix, Mol. Biol. Evol. 22 (2005) 193–199, https://doi.org/10.1093/molbev/msi005 Epub 2004 Oct 13.
[49] S. Whelan, N. Goldman, A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach, Mol. Biol. Evol. 18 (2001) 691–699, https://doi.org/10.1093/oxfordjournals.molbev.a003851.
[50] C. von Mering, M. Huynen, D. Jaeggi, S. Schmidt, P. Bork, B. Snel, STRING: a database of predicted functional associations between proteins, Nucleic Acids Res. 31 (2003) 258–261.