International Immunology, Vol. 5, No. 2, pp. 233–238 *© 1993 Oxford University Press*

A 3-D model for the CD40 ligand predicts that it is a compact trimer similar to the tumor necrosis factors

Manuel C. Peitsch and **C. Victor Jongeneel**[1]

Institute of Biochemistry, University of Lausanne and [1]Ludwig Institute for Cancer Research, Lausanne Branch, chemin des Boveresses 155, CH-1066 Epalinges, Switzerland

Key words: antigen activation, antigen differentiation, B lymphocytes, molecular models, T lymphocytes, tumor necrosis factor

Abstract

Based on the similarity in primary structure between the newly characterized ligand for CD40 (CD40L) and the tumor necrosis factors (TNFs), we have modeled a detailed 3-D structure for CD40L. We used the known structure of TNFα as a template for the generation of the CD40L model. The soundness of the model-building algorithms was verified by constructing a 3-D model of TNFβ and comparing it to its crystallographically determined structure. The CD40L sequence is entirely compatible with the 'jelly-roll' β-strand structure characteristic of the TNFs. Like the TNFs, CD40L is predicted to form a compact trimer, although the interactions between monomers are distinct from those found in the TNFs. The model predicts which regions of CD40L could interact with its receptor(s) and which amino acids are essential for the maintenance of its trimeric structure.

CD40 is a transmembrane glycoprotein expressed by B lymphocytes, but also detected on other cell types or tumors, particularly carcinomas (1). The functional characterization of CD40 has been limited to B lymphocytes, where it seems to play a key role in T cell-mediated and antigen-driven activation. Immobilized antibodies against CD40 have a number of striking effects on B cells: in synergy with agents that engage the antigen receptor and IL-4, they can induce long-term proliferation and homotypic adhesion, prevent the normal apoptotic death which occurs in germinal centers, induce the transcription of germline Ig constant regions, and activate class switching, particularly to IgE (2–6). To a large extent, these effects are similar to those induced by activated T helper cells.

CD40 is a member of a growing family of receptors having a characteristic pattern of cysteine-rich repeats in their extracellular domains (7). Other members of the family incude the low-affinity nerve growth factor (NGF) receptor, the two receptors for the tumor necrosis factors (TNFα and TNFβ), and several membrane-anchored proteins whose ligands remain to be identified (OX40, 4–1BB, CD27, and the *fas* antigen). Pox viruses also encode soluble proteins related to the type II TNF receptor, which have been shown to bind TNF (8,9). Recently, Armitage and colleagues cloned the CD40 ligand (CD40L) using a CD40/IgFc recombinant hybrid as a probe (10). They showed that CD40L is a class II transmembrane protein, specifically expressed on activated $CD4^+$ T lymphocytes, which efficiently mimics the effects of anti-CD40 antibodies when presented by transfected non-T cells. These results strongly suggest that CD40L – CD40 interactions play an important role in the T helper function, which requires direct T cell – B cell contacts.

Although the fact that CD40L could be related to TNFα and TNFβ was not immediately realized (10), the similarity between the three proteins has recently become apparent (11). When we used the BLAST server at the National Center for Biotechnology Information to search for similarities between the CD40L sequence and those deposited in the protein databases (12), we found the group of TNFαs to be the most similar at the primary sequence level, followed by the TNFβs. To explore this relationship in more detail, we performed multiple pairwise alignments using the *lalign* and *lfasta* algorithms (13). The alignment we generated (Fig. 1) differs from that published by Farrah and Smith (11) in the positions of the gaps, which were inserted so as to maximize the similarity in tertiary structure. The degrees of identity of murine CD40L with murine and human TNFα are 23.2 and 25.7% respectively, and with TNFβ 20.0 and 24.3%. This compares to a 30.6% identity between human TNFα and TNFβ, whose 3-D structures display only a 0.61 Å r.m.s. positional deviation (14). A high degree of structural conservation with a low similarity at the primary sequence level is not without precedent among proteins with a high β-strand content: among the α_2-microglobulins, the sequences of the bilin- and

Correspondence to: M. C. Peitsch

Transmitting editor: K. Arai

Received 13 September 1992, *accepted* 11 November 1992

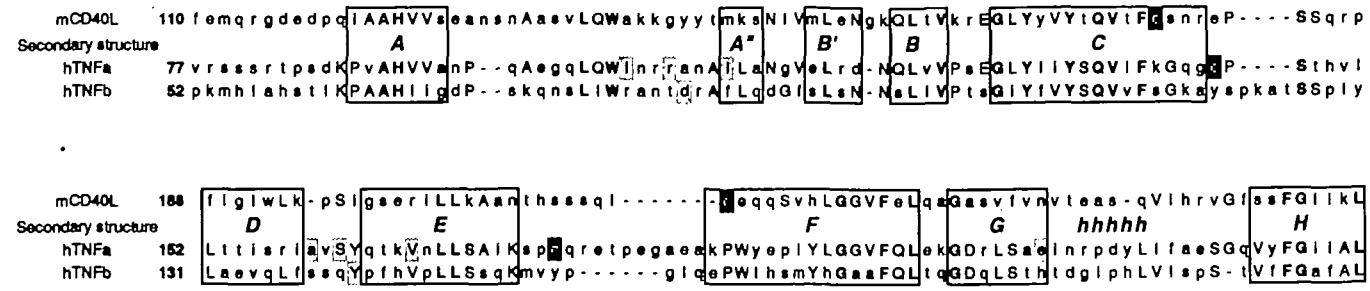

Fig. 1. Alignment of the primary sequence of CD40L with TNFα and TNFβ. The alignment was optimized to be congruent with the superposition of 3-D structures. Amino acid numbering is relative to the initiation codon, and corresponds to SwissProt database conventions. The alignments starts at the N-terminus of the secreted form of TNFα. Residues conserved between two proteins are capitalized; those conserved between all three are in boldface. Shading marks the positions in TNFα and TNFβ that are known to affect receptor binding (25,26). The cysteines participating in disulfide bonds are marked in white on black. The β-strands are boxed and labeled according to Eck *et al.* (14,18). The short α-helix in the G – H connecting region is also marked.

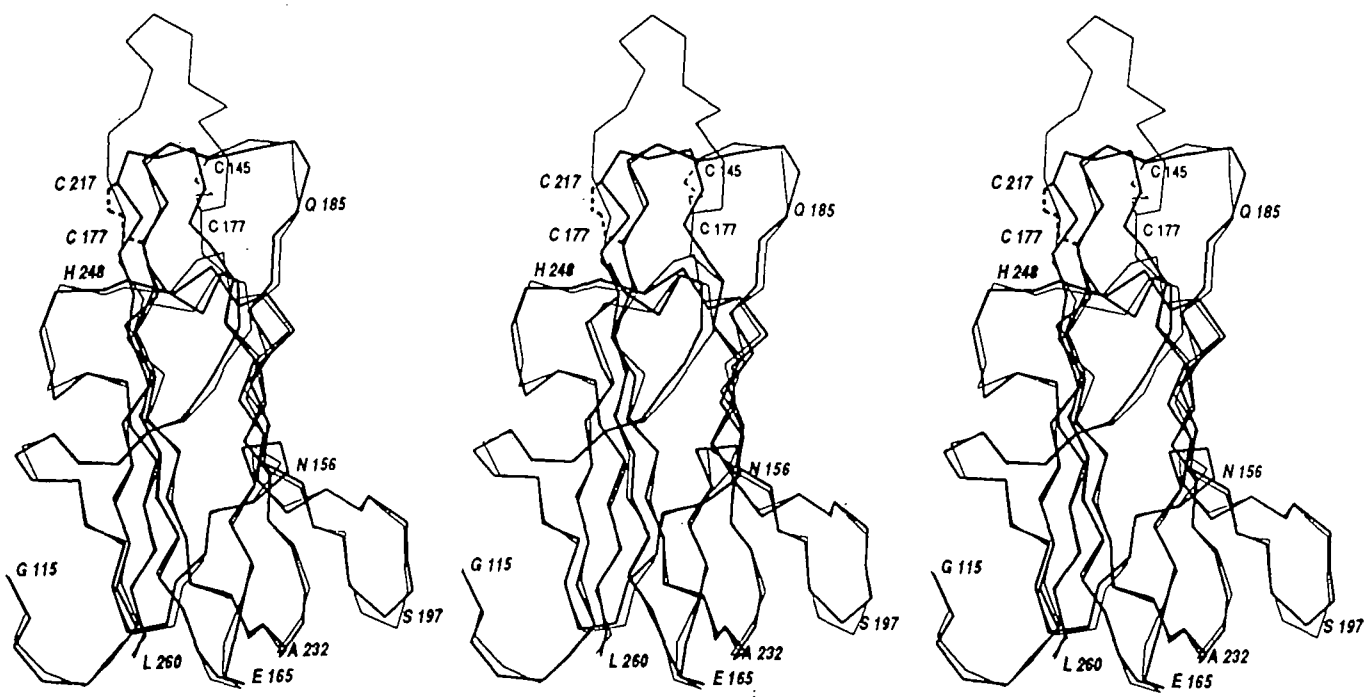

Fig. 2. Stereo representation of the superimposed α-carbon backbones of CD40L (thick line) and TNFα (thin line). Boldface amino acid positions are from the CD40L sequence. The disulfide bonds in CD40L (bold) and TNFα (thin) are marked.

retinol-binding proteins are only 16% identical, yet they differ by <2 Å r.m.s. in the α-carbon backbone of their 3-D structures (15,16). Therefore, it seemed reasonable to attempt to build a 3-D model for CD40L based on the known structure of TNFα.

As a tool to this end, we used the ProMod knowledge-based molecular modeling package [M. C. Peitsch, manuscript in preparation; ProMod is freely available from the authors (E-mail: peitsch@ulbiol.unil.ch), and has been implemented on various Unix platforms as well as under VMS], which integrates all of the following model-building steps except the energy minimization calculations: (i) construction of an averaged framework (17) from the three monomers in the TNFα trimer [entry 1TNF in the Brookhaven Protein Databank (PDB)] (18,19). (ii) Fitting of the CD40L monomer backbone onto this framework using a primary sequence alignment optimized for 3-D similarity (Fig. 1). (iii) Reconstruction of loop regions from their 'stems' by structural homology searches through the PDB databank, as described by Greer (20). (iv) Rebuilding of missing side chains using a library of allowed rotamers (21). (v) Optimization of bond geometries and relief of unfavorable non-bonded contacts by 30 steps of steepest descent followed by 500 steps of conjugate gradient minimization using the CHARMM program (22) with the PARAM19 parameter set. (vi) Superimposition of three copies of the optimized monomer structure onto the TNFα trimer, followed by another 200 steps of conjugate gradient minimization. (vii) Analysis of the resulting monomer and trimer structures using the 3-D profile matching procedure of Lüthy *et al.* (23).

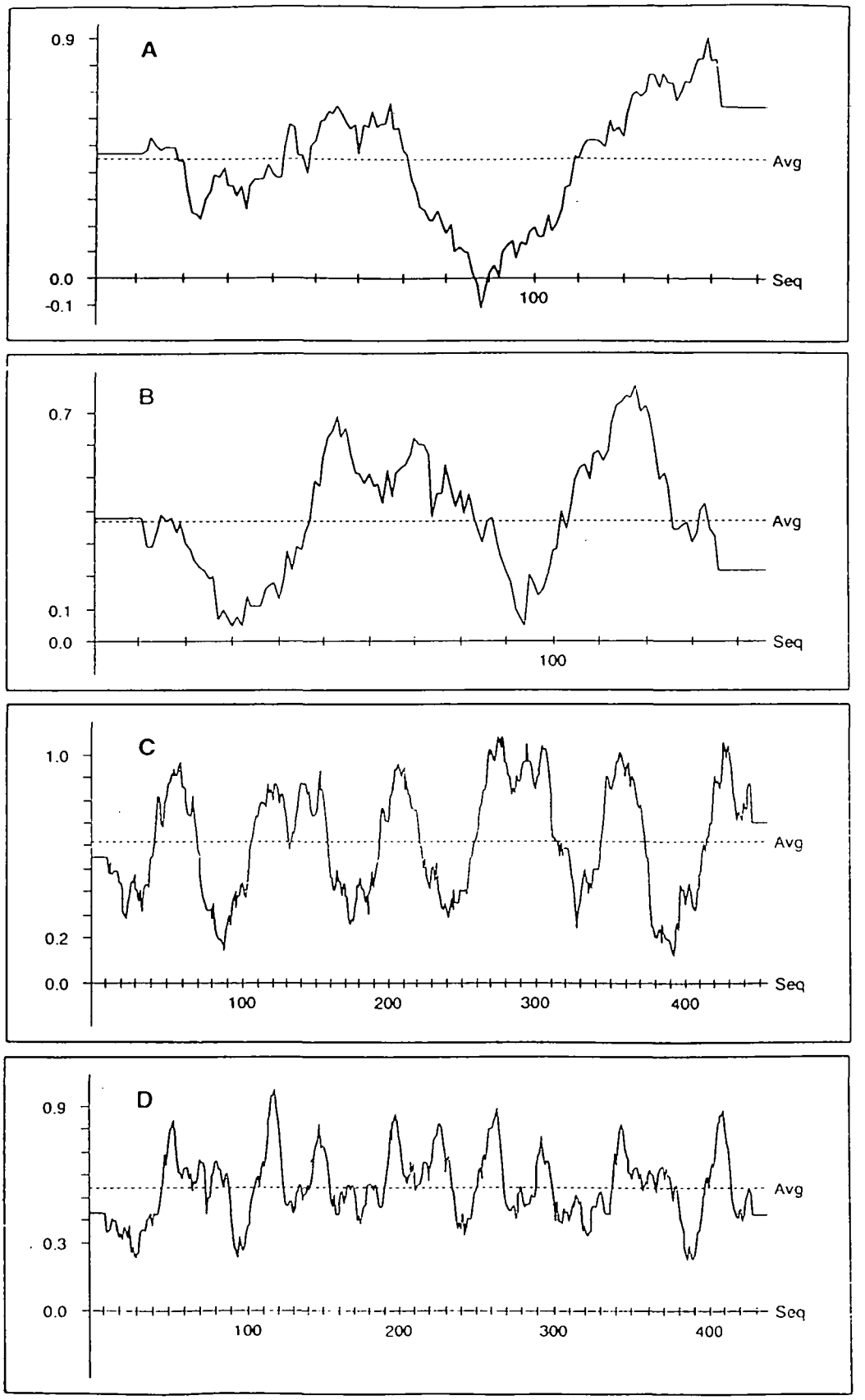

Fig. 3. Profile window plots for the TNFα and CD40L monomers (A and B) and trimers (C and D). The plots were calculated using a sliding window size of 21 to average the 3-D – 1-D scores of individual amino acids. The average score is indicated by a stippled horizontal line.

The final calculated CD40L monomer structure (coordinates of which have been deposited with the Brookhaven Protein Database) deviates from its non-minimized form by 0.52 Å r.m.s. and from the TNFα α-carbon backbone by 0.80 Å r.m.s. (Fig. 2). Verification of the monomer structure using the 3-D profile procedure (23) reveals two regions with scores falling below 0.1, which reflect unfavorably exposed non-polar surfaces (Fig. 3B). One such region also exists in the TNFα monomer structure (Fig. 3A), but none appear in the profiles of the TNFα or CD40L trimers (Fig. 3C and D). Upon trimerization, the apolar surfaces that appear as low-scoring areas in the monomer profiles participate in the contacts between subunits or are buried in regions from which water is excluded. In order to verify the validity of our modeling procedures, we constructed a 3-D model of TNFβ from the alignment shown in Fig. 1, using the same procedures as for CD40L. The modeled TNFβ trimer differed from its crystallographically determined structure (kindly provided by S. R. Sprang) by 0.95 Å r.m.s., with most of the differences confined to the unconstrained loops connecting the β strands. These results are indicative of a good-quality model with no major folding incompatibilities, and are strong evidence that CD40L exists as a native trimer, similar to TNFα and TNFβ.

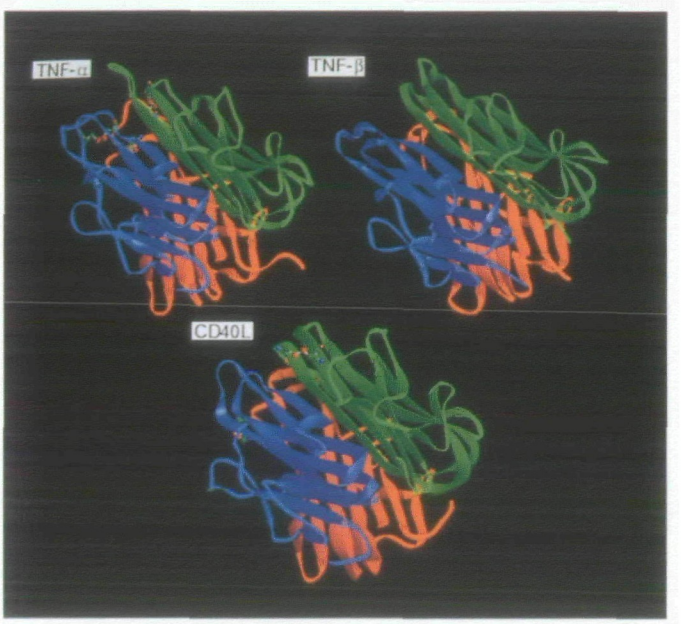

Fig. 4. Tertiary structures of the TNFα, TNFβ, and CD40L trimers. The Ribbons program was used to display the arrangement of β-sheets (flat ribbon-like regions) and the loops connecting them in the three molecules. The disulfide bonds that stabilize the TNFα and CD40L structures but are absent from TNFβ are also depicted. The pictures of TNFα and TNFβ are based on their crystallographically determined structures (14,18).

Although it is generally similar to the structure of TNFα and adopts the same 'jelly-roll' connectivity between β-strand regions (18,19), the modeled CD40L structure differs from it in several important ways (Figs 2 and 4). First, the loop connecting the E and F strands is shorter by seven residues, similar to the structure of TNFβ (14). As in TNFβ, this is accompanied by the disappearance of the disulfide bond (TNFα: Cys145 – Cys177; note that all residue numbering is relative to the initiation codon and not the N-terminus of the 'mature' form) that holds together the E – F and C – D loops in TNFα. However, another disulfide bond which directly links the C and F strands (CD40L: Cys177 – Cys217) stabilizes this region in CD40L. Compared to TNFα, the A – A′ loop is longer by two residues, while the D – E loop and the G – H linkers are shorter by one, without any major effects on the structure. The C – D loop, which is longer in TNFβ, is conserved between CD40L and TNFα. The D – E loop regions of TNFα and TNFβ are exposed to the outside and contain several residues known to influence receptor binding (Figs 1 and 5). This region is also exposed in CD40L and therefore accessible to contacts with the receptor (see below).

As mentioned above, CD40L almost certainly exists as a trimer in its native form. An examination of the monomer contact areas in our CD40L trimer model reveals some interactions which differ from the aromatic tiling and hydrophobic interactions predominating in the contact surfaces between the TNFα or TNFβ monomers (14,18,19). The Tyr169 residues are of particular interest, being in a perfect conformation to form three hydrogen bonds bracketing the 3-fold axis (Fig. 6). Just below these tyrosines, three histidines which are absent from TNFα or TNFβ also form close contacts.

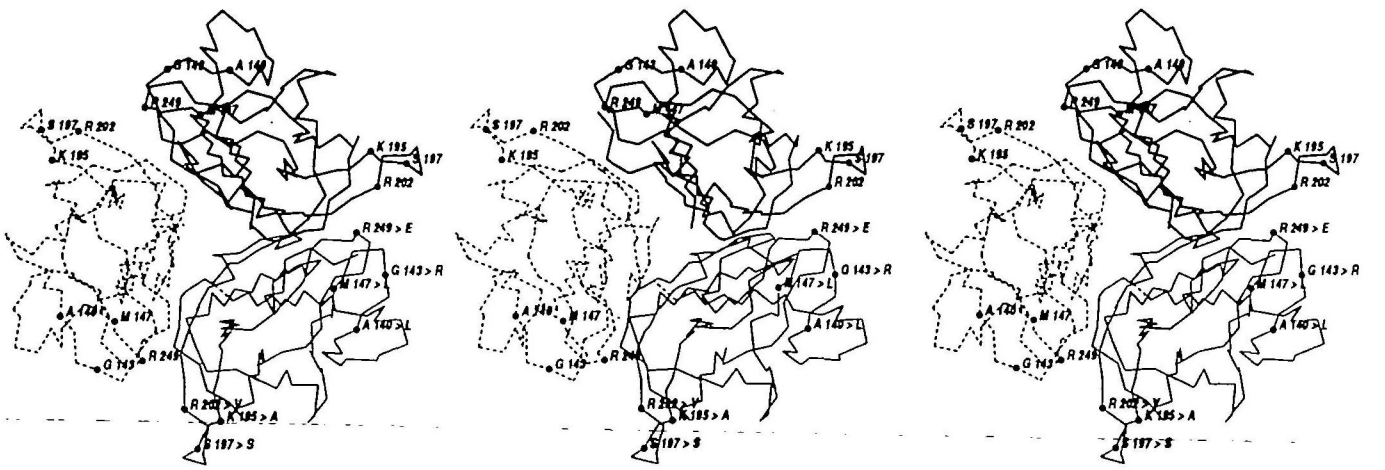

Fig. 5. Stereo view of the CD40L trimer α-carbon backbone down the 3-fold axis. Amino acids homologous to those that influence TNFα receptor binding are marked with black circles, and the corresponding residues in TNFα are indicated.

Other hydrophobic interactions are similar to those found in TNFα and β. The bottom part of the trimer, as in the TNFs, is stabilized by multiple polar interactions and salt bridges.

TNFβ recently has been shown to dock into the type II TNF receptor with the tip of its conical structure facing the cell membrane. This leads to the prediction that in TNFα and CD40L, both of which are type II membrane-anchored proteins (10,24), the N-terminal part of the trimer would face towards the base of the cone, which is proximal to the cell membrane. This prediction is borne out both in the crystal structures of TNFα and TNFβ (14,18,19) and in our model for CD40L. The specificity of the interactions between the TNFs and their receptors has been studied in some detail by random and directed mutagenesis of the ligands (25,26). In both TNFα and TNFβ, this approach has localized receptor contacts in the lower portion of the cone, in an area covering a contact surface between monomers. Surprisingly, there is very little residue conservation between TNFα and TNFβ in this area, making it difficult to predict relevant ligand–receptor contacts. The homologous region in CD40L is again very different from the TNFs': only one TNFα residue important for receptor binding is conserved in CD40L (Figs 1 and 5). It is worth noting that the necessity of cross-linking or immobilizing anti-CD40 antibodies to give them biological activity led to the speculation that the CD40 ligand should be at least a dimer. This is amply borne out by our model.

The main conclusion to be drawn from the model is that CD40L is no more different from TNFα or TNFβ than they are from each other. Similarly, the types I and II TNF receptors (TNF-RI and TNF-RII) do not resemble each other more than they do CD40, at least at the primary sequence level. In fact, construction of a similarity tree of the TNF receptor family using the PileUp program of the GCG package (27) indicates that CD40 is more closely related to TNF-RII than TNF-RI is to TNF-RII. Thus, we can identify a family of three receptors (TNF-RI, TNF-RII, and CD40) and three ligands (TNFα, TNFβ, and CD40L) which do not show any obvious structural constraints as to possible mutual interactions. The number of combinations may very well grow with the identification of the ligands for *fas*, OX40, CD27, 4–1BB, and the low-affinity NGF receptor, as NGF does not seem to be the natural ligand for this molecule (28). Chemical cross-linking experiments using TNFα as a probe have given no evidence that it binds with high affinity to surface proteins other than the two known TNF receptors (29). However, recent experiments (30) have shown that among 15 cytokines tested, TNFα was the only one able to enhance IL-4-induced synthesis of germline ε transcripts, a characteristic activity of CD40L. Therefore, it is not totally unreasonable to speculate that the TNFs and CD40L could interact with each others' receptors under some circumstances.

Finally, it is worth pointing out that, although the receptor-binding domains of TNFα, TNFβ and CD40L are extremely similar in structure, the N-terminal domains of the proteins are quite different. TNFβ is a classical secreted protein containing a signal peptide, and no transmembrane form of it has ever been detected. TNFα, on the other hand, is synthesized as a type II transmembrane protein (24) and then undergoes post-translational cleavage, liberating the extracellular domain (i.e. the trimer) as a soluble effector. Nevertheless, a substantial amount of TNFα remains membrane-associated. CD40L seems to be at the other end of the spectrum, in that is has not yet been detected as a naturally occurring secretory product (10). Ironically, it may function as a receptor for one of the many known soluble forms of members of the TNF receptor family (7). The multiple interactions between TNF-like molecules and their increasingly diverse receptors should prove to be a rich source of surprises and complexities for years to come.

Acknowledgements

We thank Dr Stephen R. Sprang, for contributing the atomic coordinates of TNFβ, and Dr Roland Lüthy, for his help with the implementation of the 3-D profile method. We gratefully acknowledge the positive criticisms and continuing support of Profs Jürg Tschopp and Jean-Charles Cerottini. We are indebted to the Computer Center of the Ecole Polytechnique Fédérale, Lausanne and the Advanced Scientific Computing Laboratory of the Frederick Cancer Research Facility for supercomputing time.

Abbreviations

NGF	nerve growth factor
PDB	Brookhaven Protein Databank
TNF	tumor necrosis factor

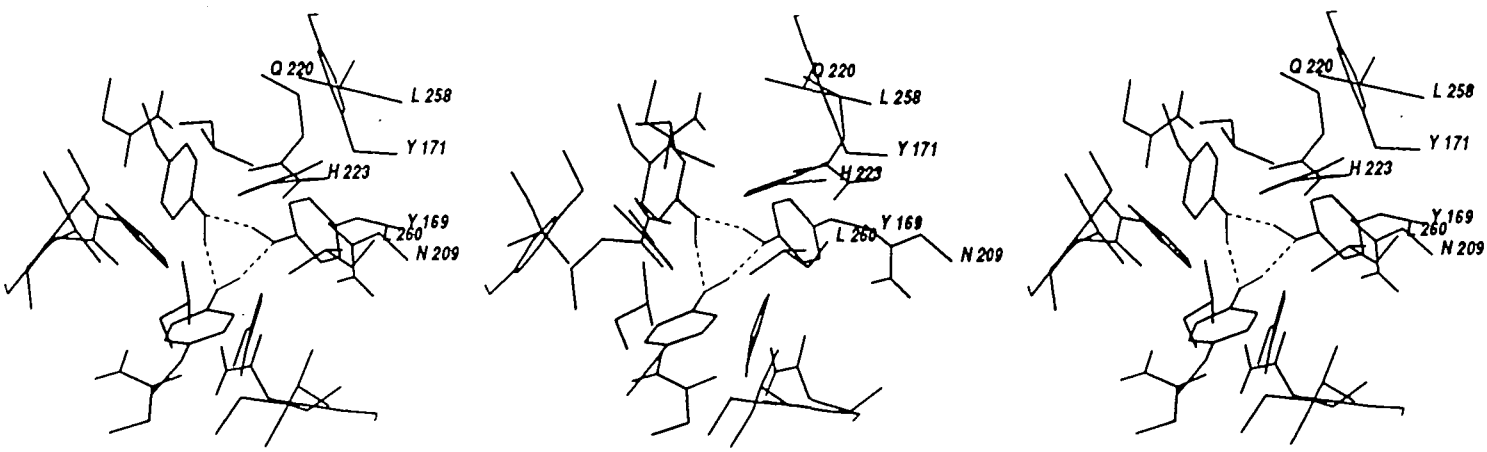

Fig. 6. Cutout stereo view of a portion of the monomer contact region. The tyrosines (Y169) forming the triple hydrogen bonds (stippled lines) are prominent in this view.

Key words There should be provided on a separate page (2) a list of key words not appearing in the title for use by indexing services.

Research papers should start with an ABSTRACT (page 3), which will appear before the main body of the text. It should be written in complete sentences and should summarize the materials, methods, results, and major conclusions in **250 words or less** in a form comprehensible to any immunologist and suitable for abstracting services. If references are used in an abstract, they must include the author(s), journal title, volume number, starting page, and year. Avoid abbreviations in abstracts as far as possible.

The main part of a research paper should start with a brief INTRODUCTION, which outlines the historical or logical origins of the study without repeating the abstract or summarizing the results, a section on METHODS, and one on RESULTS in which the observations are presented with minimal reference to earlier literature or to possible interpretations. The final DISCUSSION may usefully start with a brief summary of the major findings, but repetition of parts of the abstract or of the results section should be avoided.

Short communications should also start with an ABSTRACT. **There should be no individual sections**.

Abbreviations An alphabetized list of abbreviations used in the text must be provided on a separate page.

Acknowledgements, references, tables, and figure legends should follow the main text in that order. There should be an indication in the main text to show the most appropriate placement of each table and illustration.

6 References

References are to be cited in the text by a number in parentheses and should be numbered in order of first appearance. An example is: 'McMichael and Gotch (12) have reported . . .'. Where there are more than two authors, the citation in the text should use the formulation '*et al.*'. An example is: 'This observation has been reported by Shackelford *et al.* (24)'.

In the list of references papers from journals should be listed thus:

24 Shackelford, D. A., Smith, A. V., and Trowbridge, I. S. 1987. Changes in gene expression induced by a phorbol diester: expression of IL2 receptor, T3, and T cell antigen receptor. *J. Immunol.* 138:613.

Books should be listed as:

12 McMichael, A. J. and Gotch, F. 1987. T-cell antigens: new and previously defined clusters. In McMichael, A. J., ed., *Leukocyte Typing III*, p. 31. Oxford University Press, Oxford.

or:

18 Bowry, T. R. 1984. *Immunology Simplified*, 2nd edn. Oxford University Press, Oxford.

The titles of journals should be abbreviated in accordance with the *World List of Scientific Periodicals*, 4th edn.

7 Illustrations

Authors should use illustrations sparingly. Each figure should be relevant to the text and figures should be presented in the order in which they are mentioned in the text. **Figures and photographs should be numbered as a single series and should have the name of the first-named author, the title of the paper, and the figure number on the back**. Special instructions relating to the final size of the illustration (e.g. do not reduce; print column width) can also be written here.

The final size of printed figures cannot exceed 179 mm wide × 195 mm deep for a full-page illustration. Illustrations planned for column width cannot be printed more than 86 mm wide. All signs and letters must be clearly produced using printed paste-on or transfer lettering. Labels should be large enough for the final reduction (**letters or numbers less than 2 mm high are not acceptable in the final print**) and should show clearly against the background. Similarly, labels should not be disproportionately large. Figures should be mounted on good quality light board with a 2.5 cm margin. Where several figures are mounted together, they should be squared accurately and separated by ~5 mm. All of the figures in such a group should have approximately the same contrast values. Where regions of particular importance can be identified it is useful to indicate these on a transparent or semi-transparent overlay. Instructions to the printer regarding required contrast values or detail will help the printer to produce the best results.

Figure legends should **explain each figure as fully as possible**, referring the reader to the text only on rare occasions in order to avoid repeating in the legends material that must be included in the text. All abbreviations used in the figures should be explained in each legend when there are only a few (about five) abbreviations used. When large numbers of abbreviations are necessary in the figures, the authors should include these in a separate list of abbreviations.

Colour figures. Colour figures will be published in the journal but **authors are required to pay the costs.**

8 Criteria for FACS data in *International Immunology*

Methods Include information pertaining to all FACS data in Methods.

Cytometer. Identify the flow cytometer and describe its set-up.

Reagents. Identify all reagents. For antibody reagents include the antigen name, the specific hybridoma, and the appropriate reference (e.g. B220/RA3-6B2). Identify which fluorophores are being used. Identify any second step reagents (e.g. Texas Red – Avidin).

Calibration. Describe the calibration of the cytometer. Include the type of particles used and the method for setting the calibration points (e.g. computer program, 'by eye'). If the machine has not been calibrated, so state.

Software. Identify the software package(s) and algorithm(s) used to generate the plots. Identify the types of contours used, e.g. probability (%), log (%). Fully describe any user selected parameters.

Gating. Describe any hard or soft gates that are common to all plots (e.g. viable lymphocytes are shown after gating on forward and side scatter and propidium iodide). Describe the quantification of specific populations. For frequency data, indicate which populations are used for both the numerator and denominator (e.g. total cells, total lymphocytes defined by scatter gates).

Figures

Axes.

Scale. Clearly mark the scale intensities. It is inadequate to simply state log or linear. Channel numbers are not acceptable.

Parameter. Clearly identify the parameter for each axis, e.g. forward scatter, CD4, propidium iodide.

Gating. Draw appropriate gates on the figures when pertinent.

Legend.

Identify the tissues analyzed and other appropriate information.

Specifics. Include specifics about the reagents (e.g. PE-B220), contours, and gating that are not apparent from Methods.