CONSERVATION OF FUNCTIONAL SITES ON INTERLEUKIN-6 AND IMPLICATIONS FOR EVOLUTION OF SIGNALLING COMPLEX ASSEMBLY AND THERAPEUTIC INTERVENTION

Vaclav Veverka*, §, Terry Baker†, Nicholas T. Redpath†, Bruce Carrington†, Frederick W. Muskett*, Richard J. Taylor‡, Alastair D. G. Lawson†, Alistair J. Henry† and Mark D. Carr*

From *Department of Biochemistry, University of Leicester, Henry Wellcome Building, Lancaster Road, Leicester, LE1 9HN, UK; †UCB Pharma, 213 Bath Road, Slough, SL1 4EN, UK and §Institute of Organic Chemistry and Biochemistry, Flemingovo nam. 2, Prague, 166 10, Czech Republic

Address correspondence to: Dr Mark Carr, Department of Biochemistry, University of Leicester, Henry Wellcome Building, Lancaster Road, Leicester, LE1 9HN, UK, Tel. +44-116-229-7075; mdc12@le.ac.uk

Capsule

**Background:** Interleukin-6 signalling requires assembly of a ternary IL-6/IL-6Rα/gp130 complex.

**Results:** Determination of the mIL-6 structure allowed detailed structural and sequence comparisons with hIL-6, predicting the primacy of sites in driving IL-6/IL-6Rα-gp130 interactions, which was confirmed by binding experiments.

**Conclusion:** Interactions between gp130 domain-1 and IL-6/IL-6Rα drive signalling complex assembly.

**Significance:** Suggests a pathway for evolution of signalling complex assembly and strategies for therapeutic targeting.

A number of secreted cytokines, such as interleukin-6 (IL-6), are attractive targets for the treatment of inflammatory diseases. We have determined the solution structure of mouse IL-6 (mIL-6) to assess the functional significance of apparent differences in the receptor interaction sites (IL-6Rα and gp130) suggested by the fairly low degree of sequence similarity with human IL-6 (hIL-6). Structure-based sequence alignment of mIL-6 and hIL-6 revealed surprising differences in the conservation of the two distinct gp130 binding sites (IIa and IIIa), which suggests a primacy for site III-mediated interactions in driving initial assembly of the IL-6/IL-6Rα/gp130 ternary complex. This is further supported by a series of direct binding experiments, which clearly demonstrate a high affinity IL-6/IL-6Rα-gp130 interaction via site III but only weak binding via site II. Collectively, our findings suggest a pathway for the evolution of the hexameric, IL-6/IL-6Rα/gp130 signalling complex and strategies for therapeutic targeting. We propose that the signalling complex originally involved specific interactions between IL-6 and IL-6Rα (site I), and between the D1 domain of gp130 and IL-6/IL-6Rα (site III), with the later inclusion of interactions between the D2 and D3 domains of gp130 and IL-6/IL-6Rα (site II) through serendipity. It seems likely that IL-6 signalling benefited from the evolution of a multi-purpose, non-specific protein interaction surface on gp130, now known as the cytokine-binding homology region (site II contact surface), which fortuitously contributes to stabilisation of the IL-6/IL-6Rα/gp130 signalling complex.

IL-6 is a pro-inflammatory cytokine, which plays key roles in both innate immunity and acquired immune responses. Impaired regulation of IL-6 production often contributes to the development of chronic inflammation. IL-6 stimulates antibody production by B cells and also the differentiation of Th17 lineage T cells, which are both crucial events in experimental models of autoimmune diseases. Predictably, overproduction of IL-6 has been implicated in the pathogenesis of several chronic inflammatory diseases such as rheumatoid arthritis (for reviews see (1-3)).

The IL-6 family of cytokines signal via the activation of JAK (Janus kinase) tyrosine kinases, leading to the activation of STAT (signal transducers and activators of transcription).
transcription factors which regulate a range of target genes (4). The initial event in IL-6 signalling involves the formation of a non-signalling complex between IL-6 and its specific receptor (IL-6Rα) which leads to recruitment of gp130 to the complex through interactions with its extracellular domains and the formation of a hexameric signalling complex (5), as illustrated in figure 1. X-ray crystallographic and cryo-electron microscopy studies have revealed the molecular organization of this complex (6,7), in which IL-6 interacts with both subunits of IL-6Rα (D2 and D3) via a conserved site (site I). Regions within both proteins of the IL-6/IL-6Rα complex are involved in binding to the cytokine-binding homology region of gp130 (CHR, fibronectin type III domains D2 and D3), with this composite interface defined as site II. Additional interactions are made via another shared site (site III) with a second molecule of gp130 through its immunoglobulin-like activation domain (D1). The three distinct interaction surfaces allow assembly of a stable, ternary signalling complex composed of two molecules of each component (figure 1).

In this communication we report the solution structure of mouse IL-6, detailed analysis of the conservation of functional sites involved in signalling complex assembly and a series of binding experiments designed to assess the relative importance of specific interaction sites in ternary complex formation. The structure obtained for mIL-6 confirmed its assignment as the ortholog of human IL-6 despite the low degree of sequence homology (40% identity), and structure based sequence alignment allowed comparison of the IL-6Rα and gp130 interaction sites on human and mouse IL-6 (6,8,9). This revealed surprising differences in the conservation of the two distinct gp130 binding sites (site IIa and site IIIa, figure 1), which suggests a greater importance for site III interactions in mediating initial ternary complex assembly. This possibility was further investigated by a series of direct binding experiments, which strongly support the primacy of site III-mediated interactions in driving ternary complex formation. The work reported suggests a mechanism for the evolution of the hexameric signalling complex and highlights specific interaction sites for therapeutic targeting.

**EXPERIMENTAL PROCEDURES**

**Protein Expression and Purification** - A mature form of mIL-6 (residues 27-211) with an additional five residues at the N-terminus (Gly-Ala-Met-Gly-Ser) was produced using a pET-21-based *E. coli* expression vector, which included an N-terminal Histag with a TEV protease cleavage site in the linker. Tuner (DE3) pLysS cells (Novagen) transformed with the mIL-6 expression vector were grown at 37°C in media containing 100 μg/ml carbenicillin to an A600 of approximately 0.8, then cooled to 25°C and mIL-6 expression induced overnight by the addition of 10 μM IPTG. Cells were harvested, lysed and the Hisc-tagged mIL-6 purified using a Ni-NTA column (Qiagen). The His-tag was subsequently removed by incubation of the purified protein with TEV protease (TEV:mIL-6 1:15 w/w) at room temperature overnight followed by a second round of Ni-NTA affinity chromatography. Fractions containing mIL-6 were subsequently dialysed into 20 mM sodium phosphate and 100 mM sodium chloride buffer at pH 6.4 and subjected to a final gel filtration purification step on a Superdex 75 16/60 column. The mature human IL-6 (hIL-6, residues 28-212) was prepared using an analogous pET-21-based *E. coli* expression vector (N-terminal His6 tag with a TEV protease cleavage site), with expression carried out in Origami B (DE3) cells induced by the addition of 100 μM IPTG and grown overnight at 17°C. The soluble hIL-6 produced was harvested and purified essentially as described for mIL-6.

Uniformly 15N and 13C labelled mIL-6 was produced from cells grown in minimal medium containing 1 g/l 15N ammonium sulphate and if appropriate 2 g/l 13C D-glucose as the sole nitrogen and carbon sources. To improve the quality of the 13C/H HSQC-NOESY spectra, 15N/13C labelled samples of mIL-6 were prepared that contained non-labelled aromatic residues (Phe, Tyr, Trp and His), which was achieved by the addition of 50 mg/l of non-labelled amino acids to the minimal medium.

The human IL-6Rα-IL-6 fusion protein was expressed with a cleavable N-terminal His6 tag in CHO cells and is similar to the hyper IL-6 fusion protein described previously (10). The soluble IL-6Rα-IL-6 was initially harvested from cell culture supernatants by Ni-NTA affinity
were removed by a second round of Ni-NTA tag and any remaining His$_6$ tag by TEV protease treatment. The released His$_6$-tagged fusion protein was expressed in CHO cells and purified as described previously (11). This briefly involved isolation of the soluble gp130-Fc or IL-6Ralpha-Fc fusion proteins from culture supernatants by Protein A affinity chromatography followed by gel filtration on a Sephacryl S200 16/60 column. The pure human oncostatin M and human gp130 extracellular region-Fc fusion protein (domains D1 to D6) were obtained from R&D Systems.

**NMR Spectroscopy** - NMR spectra were acquired from 0.35 ml samples of mIL-6 (250-430 μM) in a 20 mM sodium phosphate and 100 mM sodium chloride buffer at pH 6.4 (95% H$_2$O/ 5% D$_2$O or 100% D$_2$O). All NMR experiments were collected at 25°C on either 600 or 800 MHz Bruker spectrometers equipped with triple-resonance ($^{15}$N/$^1$H/$^2$H) cryoprobes. A series of double and triple-resonance spectra were recorded to determine essentially complete sequence-specific resonance assignments for mIL-6, as described previously (12-14). $^1$H-$^1$H distance constraints required to calculate the structure of mIL-6 were derived from NOEs identified in NOESY, $^{15}$N/$^1$H NOESY-HSQC and $^{13}$C/$^1$H HMQC-NOESY spectra, which were acquired with an NOE mixing time of 100 ms. Residues involved in stable backbone hydrogen bonds were identified by monitoring the rate of backbone amide exchange in 2D $^{15}$N/$^1$H HSQC spectra of mIL-6 dissolved in D$_2$O.

**Structural Calculations** - The family of converged mIL-6 structures was calculated using CYANA 2.1(15). Initially, the combined automated NOE assignment and structure determination protocol (CANDID) was used to automatically assign the NOE cross-peaks identified in 2D NOESY and 3D $^{15}$N- and $^{13}$C-edited NOESY spectra, and to produce preliminary structures. Subsequently, several cycles of simulated annealing combined with redundant dihedral angle constraints (REDAC) were used to produce the final converged mIL-6 structures (16). Analysis of the family of converged structures obtained was carried out using the programs CYANA, MOLMOL, iCING, and PyMol (15,17-19).

**Protein Sequence Alignments** - The sequence alignments for human and mouse IL-6, IL-6Ralpha and gp130 were generated using ClustalW (20). Structure-based sequence alignments of mIL-6 with both free and signalling complex associated hIL-6 (IL-6/IL-6Ralpha/gp130, PDB accession codes 1IL6 and 1P9M respectively) were obtained by detailed comparisons of the structures in PyMol and sequence analysis in JalView (21).

**Surface Plasmon Resonance** - All surface plasmon resonance (SPR) experiments were carried out on a Biacore 3000 system using a pH 7.4 running buffer containing 10 mM HEPES, 150 mM NaCl and 0.005% (v/v) P20. Proteins were attached to the surface of CM5 sensor chips (GE Healthcare) by the amine coupling method, as recommended by the manufacturers. Briefly, the carboxymethyl dextran surface was activated with a fresh mixture of 50 mM N-hydroxysuccimide and 200 mM 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide for 5 minutes at a flow rate of 10 μl/min. The gp130-Fc and IL-6Ralpha-IL-6 fusion proteins were covalently coupled to chips at 7.5 μg/ml and 2 μg/ml respectively in a 10 mM acetate pH 5.0 buffer using a 5 minute pulse at the same flow rate. Finally the surface was deactivated with a 10 minute pulse of 1 M ethanolamine.HCl pH8.5. Reference flow cells were prepared for each type of chip produced by omitting the protein from the above procedure.

**Binding of Native and Single Residue Variants of gp130-D1 to IL-6Ralpha-IL-6 Fusion** - A series of purified native and variant gp130-D1-Fc fusion proteins (E23A, L24A and L25A) were diluted in running buffer over the range 25 to 250 mM and pulsed over reference and IL-6Ralpha-IL-6 fusion flow cells for 4 minutes followed by a 5 minute dissociation phase in running buffer. The chip surfaces were regenerated with a 40 second pulse of 2 M guanidine.HCl at 30 μl/min. Sensorgrams were obtained as the response unit difference between the immobilised protein and reference cell.
Relative Binding of Functional Partners and Complexes to Extracellular Regions of gp130 -Solutions of hIL-6, IL-6RA-D1D2D3-Fc, oncostatin M and an equimolar mixture of hIL-6 and IL-6RA-D1D2D3-Fc were prepared in SPR running buffer at a concentration of 50 nM. The individual samples were each pulsed at 20 μl/min over immobilized gp130-D1-Fc, gp130-D2D3-Fc or gp130-D1D2D3D4D5D6-Fc fusion proteins for 150 seconds and a binding response report point was taken 10 seconds into the dissociation phase. The chip surfaces were regenerated with a 60 second pulse of 2.5 M guanidine.HCl at 30 μl/min.

The relative binding of functional partners and complexes to immobilized gp130-Fc fusion proteins was determined by normalising the report point responses obtained by the molecular weight ratios for the appropriate binding partners to gp130-Fc constructs, and by correcting for any differences in the efficiency of the coupling of gp130 proteins to sensor chips.

RESULTS

Solution Structure of mIL-6. Comprehensive sequence-specific backbone and side-chain resonance assignments were obtained for mIL-6 using a proven combination of triple-resonance experiments, as described previously (22). Backbone amide signals ($^{15}$N and $^1$H) were assigned for all residues apart from Thr27-Ser28, Gln94-Gly116 and Lys177-Thr189, which were not detectable, probably due to line-broadening arising from conformational exchange processes in the affected regions. Assignments obtained for non-exchangeable side-chain signals were over 84.5% complete (nearly 95% complete for regions of the protein with assigned backbone resonances).

The completeness of the $^{13}$N, $^{13}$C and $^1$H resonance assignments obtained for mIL-6 allowed automated assignment of the NOEs identified in 3D $^{15}$N/$^1$H NOESY-HSQC, $^{13}$C/$^1$H HMBC-NOESY and in the aromatic to aliphatic region of 2D NOESY spectra using the CANDID procedure implemented in CYANA (15). This approach proved very successful and yielded unique assignments for 94.2% (3863/4103) of the NOE peaks observed, which provided 2914 non-redundant $^1$H-$^1$H distance constraints. In the final round of structural calculations, 52 satisfactorily converged mIL-6 structures were obtained from 100 random starting conformations using a total of 3418 NMR-derived structural constraints (an average of 18 constraints per residue). Table 1 contains a detailed summary of the constraints included in the calculations, together with the structural statistics for the family of structures obtained. The mIL-6 structures, together with the NMR constraints, have been deposited in the Protein Data Bank under accession number 2L3Y.

The solution structure of mIL-6 has been determined to relatively high precision, which is clearly evident from the overlay of the protein backbone shown for the family of converged structures in figure 2A. This is also reflected in reasonably low root mean squared deviation (rmsd) values to the mean structure for both the backbone (0.64 ± 0.13 Å) and all heavy atoms (1.26 ± 0.13 Å, Table 1) of residues forming the well defined core of mIL-6 (Val45-Lys89, Ile113-Leu160 and Arg186-Thr208).

mIL-6 contains four relatively long α-helices (helix A: Thr48-Leu69, helix B: Ser114-Asn129, helix C: Asn134-Asp159 and helix D: Arg186-Thr208), which form a four-helix bundle that is decorated by two shorter and less well defined α-helices (helix S1: Asp75-Asn86 and helix S2: Ile169-Asp176) and by several loops, as shown in figure 2B. The 25 N-terminal residues of the protein appear to be highly flexible and unstructured on the basis of the NMR data. The irregular loops connecting helical regions are a major feature of mIL-6, but differ widely in their length and properties. The short loop between helix A and helix S1 is fairly well defined, which probably reflects the presence of a disulphide bond between Cys70 and Cys76. This region is further stabilised by a network of favourable van der Waals contacts between helix S1 and helix D. The long loop between helix S1 and helix B is poorly defined by the NMR data and corresponds to a region in which many backbone amide signals are missing, which is consistent with the existence of multiple conformations in intermediate exchange on the NMR timescale. In contrast, the long loop between helices C and D, which includes the short S2 helix, is well defined apart from the region between helices S2 and D. These variable properties for the loop regions are reflected in the family of mIL-6 structures shown in figure 2A.

Comparison of the Structures of Mouse and Human IL-6. The predicted mouse ortholog of
human IL-6 shares only 40% amino-acid sequence identity and about 52% overall sequence homology (23,24), which is surprisingly low compared to all the other cytokines that signal through the gp130 receptor (interleukin-11 (88/92%), leukemia inhibitory factor (78/87%), ciliary neurotrophic factor (82/86%), cardiotrophin-1 (80/86%), cardiotrophin-like cytokine (97/99%) and interleukin-27 (69/73%)), apart from oncostatin M (42/50%). The remaining class-I helical cytokines (25), which signal through a gp130-independent mechanism, also show a much higher degree of sequence homology (supplementary figure 1).

Determination of the solution structure of mIL-6 has allowed detailed comparisons with the previously reported structures for free and ternary signalling complex bound hIL-6 (6,8,9). The human and mouse proteins are both characterised by a four helix bundle at the heart of the structure, and despite the relatively low sequence homology the positions and orientations of the four long α-helices in mIL-6 and hIL-6 are fairly well conserved, which is reflected in an rmsd of 1.8 Å for equivalent backbone atoms superimposed (figure 2C). The angles between pairs of helices differ by approximately 16° for A compared to B and C. In addition, helices A and C are somewhat longer in mIL-6 (3 and 7 residues respectively), whereas helices B and D are significantly shorter (7 and 5 residues respectively). Helix S1 is formed by 13 residues in mIL-6 and in contrast to either solution or crystal structures of hIL-6 is well defined (8,9), which implies that this region is more stable in mIL-6, perhaps reflecting a more extensive contact interface with the helical bundle.

In the human hexameric signalling complex (figure 1), helix S1 is shared between sites I (IL-6/IL-6Rα) and IIIa (IL-6/gp130-D1) (6). The greater stability of this helix in mIL-6 may lead to higher affinity binding at these sites due to a lower entropic penalty. Interestingly, additional mobility seen in the hIL-6 S1 helix (8,9) is maintained in the hexameric signalling complex (6). The loop between helices S2 and D is surprisingly similar in mIL-6 and hIL-6 (8,9), which may reflect its conserved functional significance as part of the site III binding interface (figure 3). It is likely that this loop undergoes a conformational change on binding to the D1 domain of gp130 similar to that seen for hIL-6. Notable differences between mIL-6 and hIL-6 occur in the loop connecting helix S1 to helix B, which is significantly shorter in the human protein and results in reduced mobility in hIL-6 (8,9). However, this region is not directly involved in signalling complex formation (6).

Conservation of Functional Sites on IL-6 and Cell Surface Receptors. A structure-based sequence alignment of mouse and human IL-6, together with optimal sequence alignments for the interacting regions of the corresponding IL-6 receptors (IL-6Rα) and co-receptors (gp130), are shown in figure 4, which reveals significant variations in the conservation of regions involved in key interactions required to form the hexameric IL-6/IL-6Rα/gp130 signalling complex (6). The first stage of signalling complex formation involves the interaction of IL-6 with its dedicated receptor (IL-6Rα) through surface regions on both proteins defined as site I (6). The conservation of residues involved in this interaction is relatively high (50–80% identity) across a representative set of mammalian species, as shown in figures 4 and 5. The next stage in signalling complex assembly is proposed to be the interaction of the IL-6/IL-6Rα complex with the D2 and D3 domains of gp130 (IL-6Rβ) (6), which involves a cluster of 12 residues localized on helices A and C of IL-6 (site IIa). The region of IL-6 involved in this interaction is very poorly conserved across mammalian species (8–33% identity, figure 5B), with only a single residue unchanged between mouse and human (figure 4). In contrast, the complementary site IIa on gp130 is well conserved (50–80%, figures 4 and 5) with no evidence of changes introduced to compensate for residue changes in IL-6. This strongly suggests that contacts at site IIa stabilise rather than drive tertiary IL-6/IL-6Rα/gp130 complex formation. The lack of conservation of site IIa interactions contrasts markedly with the other interface involved in binding of IL-6 to gp130 (site IIIa), which is the most conserved on the protein (68–74%), as illustrated in figures 4 and 5C. This interface is mediated by the immunoglobulin-like D1 domain of gp130 and residues on IL-6 present in the flexible loop between helix S2 and helix D and also in helix S1 (figure 3).

Roles of Specific Functional Sites in Mediating Binding of gp130 to IL-6/IL-6Rα Complex. To
directly assess the role of site IIIa in driving signalling complex formation, a series of SPR experiments were carried out to determine the affinity of the interaction between immobilised IL-6Rα-IL-6 fusion and native as well as variant gp130-D1-Fc fusion proteins (E23A, L24A and L25A). The typical sensorgrams shown in figure 6 clearly demonstrate a high affinity interaction ($K_d \sim 25-75$ nM) mediated by the D1 domain of gp130, which is dramatically reduced or abolished by just single residue substitutions of conserved leucine residues (L24A and L25A) within the site IIIa region of gp130 (figure 4). Although highly conserved, these N-terminal residues play no role in stabilising the structure of gp130-D1 (6), so the dramatic affects on binding resulting from the L24A and L25A substitutions almost certainly reflect the loss of key interactions with IL-6Rα-IL-6 rather than disruption of the structure of gp130-D1. The specificity of this interaction is further confirmed by the lack of any significant affect on binding of substitution of the preceding residue (E23A), which lies outside of the site IIIa interface.

SPR measurements were also used to determine the relative affinities for binding of functional partners and complexes to selected extracellular regions of immobilised gp130 (D1, D2D3 and D1D2D3D4D5D6), and normalised dissociation phase report point responses from a typical series of experiments are shown in figure 7. The results clearly indicate that neither IL-6 nor IL-6Rα alone show any stable interaction with the extracellular region of gp130, as expected from previously reported studies (6). In contrast, oncostatin M clearly forms a stable interaction with both the D1D2D3D4D5D6 and D2D3 regions of gp130, which is consistent with previous work (6) and confirms the functional integrity of the gp130-Fc constructs. As expected, the IL-6/IL-6Rα complex shows a strong interaction with the complete extracellular region of gp130, however, only a fairly weak interaction is seen with the D2D3 region of gp130 in which the binding interface is restricted to site II. This relatively weak binding between IL-6/IL-6Rα and gp130-D2D3 contrasts markedly with the strong binding observed between the complex and gp130-D1 (figure 7), which clearly points to interactions mediated via site III as the principal drivers of ternary IL-6/IL-6Rα/gp130 complex formation. Collectively, the SPR results reported here, together with consideration of the sequence conservation of the distinct interaction sites (discussed above), strongly support the primacy of site III-mediated binding in driving ternary complex formation, with site II-mediated interactions facilitating the subsequent formation of the hexameric signalling complex. Previous work has also implied a critical role for gp130-D1 in ternary IL-6/IL-6Rα/gp130 complex formation, with the introduction of an N-terminal FLAG tag preceding P27 of gp130 shown to severely impair the ability of the protein to interact with IL-6/IL-6Rα complex (26), which presumably reflects the loss of several conserved site IIIa residues (L24, L25 and D26, figure 4).

**DISCUSSION**

The region of gp130 involved in site IIA interactions is known as the cytokine homology region and forms the binding site for a diverse range of cytokines utilizing gp130 as their co-receptor including IL-6, IL-11, LIF and CNTF (figure 5 25,27,28). The features of this interaction site on gp130 reflect its broad specificity, with a fairly small ($600 \text{ Å}^2$) and shallow contact surface. In contrast, there is an extensive contact surface between gp130 and IL-6Rα at site IIB ($1100 \text{ Å}^2$), which appears to contribute the majority of the specificity and affinity of interactions at site II (6). The second interface involved in binding of IL-6 to gp130 (site IIIa) is similarly rich in surface features and specific interactions between the proteins, as illustrated in figure 3, with contacts mediated by residues within domain 1 of gp130 and present in the flexible loop between helices S2 and D of IL-6 and in helix S1 (6). Interestingly, the surface of gp130 domain D1 involved in binding to IL-6/IL-6Rα is shared by IL-11, which appears to form a similar hexameric signalling complex (29,30). This site is also used by a functional homolog of hIL-6 from Kaposi’s sarcoma-associated herpesvirus (vIL-6), which in contrast to IL-6 has no requirement to interact with IL-6Rα and signals via formation of a tetrameric vIL-6/gp130 complex (31,32). It has been proposed previously that interactions between IL-6/IL-6Rα and gp130 at site II (figure 1) facilitate the initial formation of the ternary IL-6/IL-6Rα/gp130 complex (6).
However, the contrasting conservation and features of sites II and III, together with directly comparative binding data indicating a significantly higher affinity interaction via site III, strongly suggest that binding via site III is likely to drive initial formation of the ternary complex, whereas site II-mediated interactions are likely to facilitate subsequent formation of the hexameric signalling complex.

At first glance the sequence divergence in the site IIa interaction site of IL-6 across mammalian species (figure 5B) is somewhat surprising, however, it is likely that IL-6 sequences have been able to diverge here because of the essentially non-specific nature of the site IIa interaction surface on gp130, which is a shared binding site for at least eight distinct cytokines signalling through gp130 (25). Our structural comparisons of the mouse and human IL-6 proteins, together with analysis of sequence conservation across mammalian species and binding experiments assessing the affinities of site II and site III mediated interactions, have clear implications in terms of evolution of the IL-6 signalling complex, and strongly suggest that the system developed through specific interactions at sites I and III followed by later utilisation of site II. It seems very probable that IL-6 signalling is simply taking advantage of the fact that site IIA of gp130 evolved into a multi-purpose, fairly non-specific protein interaction surface, which has fortuitously been enrolled in stabilisation of a super-signalling IL-6/IL-6Rα/gp130 complex.

The known biological roles of IL-6 make it an attractive target for therapeutic intervention in a number of major inflammatory diseases, such as rheumatoid arthritis, and its extracellular location points to therapeutic antibodies as a viable approach. Clearly, consideration of the relative importance of specific protein interaction sites in signalling complex assembly and stabilisation is an important factor in selecting antibodies as therapeutic candidates, and highlights the value of detailed characterisation of the properties and features of functional sites on target therapeutic proteins and complexes.

References

1. Jones, S. A. (2005) *Journal of immunology* **175**, 3463-3468
2. McInnes, I. B., and Schett, G. (2007) *Nature Reviews Immunology* **7**, 429-442
3. Finckh, A., and Gabay, C. (2008) *Current Opinion in Rheumatology* **20**, 269-275
4. Heinrich, P. C., Behrmann, I., Haan, S., Hermans, H. M., Muller-Newen, G., and Schaper, F. (2003) *Biochemical Journal* **374**, 1-20
5. Taga, T., and Kishimoto, T. (1997) *Annual Review of Immunology* **15**, 797-819
6. Boulanger, M. J., Chow, D. C., Brevnova, E. E., and Garcia, K. C. (2003) *Science* **300**, 2101-2104
7. Skiniotis, G., Boulanger, M. J., Garcia, K. C., and Walz, T. (2005) *Nature Structural & Molecular Biology* **12**, 545-551
8. Xu, G. Y., Yu, H. A., Hong, J., Stahl, M., McDonagh, T., Kay, L. E., and Cumming, D. A. (1997) *J Mol Biol* **268**, 468-481
9. Somers, W., Stahl, M., and Seehra, J. S. (1997) *EMBO J* **16**, 989-997
10. Fischer, M., Goldschmitt, J., Peschel, C., Brakenhoff, J. P., Kallen, K. J., Wollmer, A., Grotzinger, J., and Rose-John, S. (1997) *Nat Biotechnol* **15**, 142-145
11. Schroers, A., Hecht, O., Kallen, K. J., Pachta, M., Rose-John, S., and Grotzinger, J. (2005) *Protein Sci* **14**, 783-790
12. Veverka, V., Lennie, G., Crabbe, T., Bird, I., Taylor, R. J., and Carr, M. D. (2006) *J Biomol NMR* **36**, 3-3
13. Waters, L. C., Bohm, M., Veverka, V., Muskett, F. W., Frenkel, T. A., Kelly, G. P., Prescott, A., Dosanjh, N. S., Kлеппнер, K. H., and Carr, M. D. (2006) *J Biomol NMR*
14. Renshaw, P. S., Veverka, V., Kelly, G., Frenkel, T. A., Williamson, R. A., Gordon, S. V., Hewinson, R. G., and Carr, M. D. (2004) *J Biomol NMR* **30**, 225-226
15. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) *J Mol Biol* **319**, 209-227
16. Guntert, P., and Wuthrich, K. (1991) *J Biomol NMR* **1**, 447-456
17. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J Mol Graph* **14**, 51-55
18. DeLano, W. L. (2002) [http://www.pymol.org](http://www.pymol.org)
19. Doreleijers, J. F., Sousa da Silva, A. W., Krieger, E., Nabuurs, S. B., Spronk, C. A., Stevens, T. J., Vranken, W. F., Vriend, G., and Vuister, G. W. (2012) *J Biol Mol NMR*
20. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) *Bioinformatics* **23**, 2947-2948
21. Clamp, M., Cuff, J., Searle, S. M., and Barton, G. J. (2004) *Bioinformatics* **20**, 426-427
22. Veverka, V., Crabbe, T., Bird, I., Lennie, G., Muskett, F. W., Taylor, R. J., and Carr, M. D. (2008) *Oncogene* **27**, 585-595
23. Simpson, R. J., Moritz, R. L., Rubira, M. R., and Vansnick, J. (1988) *European Journal of Biochemistry* **176**, 187-197
24. Tanabe, O., Akira, S., Kamiya, T., Wong, G. G., Hirano, T., and Kishimoto, T. (1988) *Journal of Immunology* **141**, 3875-3881
25. Wang, X., Lupardus, P., Laporte, S. L., and Garcia, K. C. (2009) *Annu Rev Immunol* **27**, 29-60
26. Moritz, R. L., Ward, L. D., Tu, G. F., Fabri, L. J., Ji, H., Yasukawa, K., and Simpson, R. J. (1999) *Growth Factors* **16**, 265-278
27. Boulanger, M. J., and Garcia, K. C. (2004) *Cell Surface Receptors* **68**, 107+
28. Skiniotis, G., Lupardus, P. J., Martick, M., Walz, T., and Garcia, K. C. (2008) *Molecular Cell* **31**, 737-748
29. Barton, V. A., Hall, M. A., Hudson, K. R., and Heath, J. K. (2000) *J. Biol. Chem.* **275**, 36197-36203
30. Matadeen, R., Hon, W. C., Heath, J. K., Jones, E. Y., and Fuller, S. (2007) *Structure* **15**, 441-448
31. Hoischen, S. H., Vollmer, P., Marz, P., Ozbek, S., Gotze, K. S., Peschel, C., Jostock, T., Geib, T., Mullerg, J., Mechtersheimer, S., Fisher, M., Grotsinger, J., Galle, P. R., and Rose-John, S. (2000) *European Journal of Biochemistry* **267**, 3604-3612
32. Chow, D. C., He, X. L., Snow, A. L., Rose-John, S., and Garcia, K. C. (2001) *Science* **291**, 2150-2155

**FOOTNOTES** - We would like to thank Laura Griffin and Shirley Peters for help with some molecular biology aspects of the reported work, and also Hanna Hailu, Robert Griffin and Mandy Oxbrow for supporting the expression and purification of proteins required for the SPR experiments.

**FIGURE LEGENDS**

**FIG. 1.** Schematic representation of the IL-6/IL-6Rα/gp130 signalling complex. The formation of a hexameric signalling complex is strongly supported by crystallographic data (6,7) and includes two molecules of each component: IL-6 (red), IL-6Rα (green; D2 and D3 domains) and gp130 (blue; D1, D2 and D3 domains). Five distinct protein interaction sites are highlighted.

**FIG. 2.** Solution structure of mIL-6. Panel A shows a best-fit superposition of the protein backbone (Val45-Thr208) for 20 out of 52 converged structures obtained for mIL-6, whilst panel B contains a ribbon representation of the backbone topology of the structure closest to the mean in the same orientation. The semi-flexible character of two loops is emphasized by arrows in panel A. Panel C shows a comparison of the backbone topologies for mouse (green) and human (pink) IL-6.

**FIG. 3.** Structural organisation of site III in the IL-6/IL-6Rα/gp130 complex. Panel A shows the features seen in the crystal structure determined for the hexameric IL-6/IL-6Rα/gp130 signalling complex (1P9M, (6)). In panel B the family of solution structures obtained for mIL-6 have been superimposed on hIL-6 to illustrate the similarity in the loop region connecting helices S2 and D, which makes key interactions between IL-6 and domain D1 of gp130 (site IIIa).
FIG. 4. Conservation of functional sites between mouse and human orthologs of the IL-6/IL-6Rα/gp130 signalling complex. The figure shows a structure-based sequence alignment (top) encompassing the well defined regions of human (residues 48-211) and mouse IL-6 (residues 46-210), together with optimised sequence alignments for the interacting regions of human and mouse IL-Rα (residues 123-303 and 119-300 respectively, middle) and gp130 (residues 23-290 and 23-288 respectively, bottom). Identical residues are shaded in dark grey and conserved residues in light grey. Arrows indicate the residues involved in specific protein interaction sites. The alignments were prepared with JalView (21).

FIG. 5. Conservation of interaction sites in class-I helical cytokines that signal through gp130. The bar charts show the sequence identity of functional sites I (A), II (B) and III (C) for a representative selection of mammalian proteins compared to the human sequences. For comparison, the conservation of the complementary sites on gp130 and IL-6Rα are also shown.

FIG. 6. Binding of gp130 domain 1 (D1) to immobilised IL-6Rα-IL-6 fusion protein. Typical SPR sensorgrams are shown for native and single residue variants of gp130-D1 binding to immobilised IL-6Rα-IL-6, with the responses obtained for the native (A) and E23A (B) proteins consistent with a relatively tight interaction (Kd ~ 25-75 nM). The two site IIIa variants of gp130-D1 (C and D) show evidence of very little if any interaction with IL-6Rα-IL-6.

FIG. 7. Relative binding of functional partners and complexes to gp130 extracellular regions. The histogram shown summarises typical normalised report point responses obtained from the dissociation phases of SPR sensorgrams with immobilised gp130 extracellular regions (D1, D2D3 and D1D2D3D4D5D6) and either IL-6, IL-6Rα, oncostatin M or IL-6/IL-6Rα complex in solution. The responses seen highlight very rapid dissociation of both IL-6 and IL-6Rα from all regions of gp130, whereas gp130 constructs containing domain 1 show slow dissociation of the IL-6/IL-6Rα complex consistent with a tight interaction via site III.
TABLE 1 - NMR constraints and structural statistics for mIL-6.

| Number of constraints used in the final structural calculations |  |
|---------------------------------|---|
| Intraresidue NOEs                | 713 |
| Sequential NOEs ($i, i+1$)      | 725 |
| Medium-range NOEs ($i, i > 1 i \leq 4$) | 978 |
| Long-range NOEs ($i, i \geq 5$)  | 498 |
| Torsion angles                   | 204 (102 $\Phi$ and 102 $\Psi$) |
| Backbone hydrogen bonds          | 300 (75 hydrogen bonds) |

| Maximum and total constraint violations in the 52 converged structures |  |
|-------------------------------------------------------------|---|
| Upper distance limits (Å)                                   | 0.18 ± 0.05 |
| van der Waals contacts (Å)                                  | 0.24 ± 0.02 |
| Torsion angle ranges ($^\circ$)                             | 2.06 ± 0.45 |
| Average CYANA target function (Å$^2$)                       | 1.13 ± 0.27 |

| Structural statistics for the family of 52 converged mIL-6 structures |  |
|---------------------------------------------------------------------|---|
| Residues within allowed regions of the Ramachandran plot            | 98.9 % |
| Backbone atom rmsd for structured regions (residues 24-68,92-139,165-187) | 0.64 ± 0.13 Å |
| Heavy atom rmsd for structured regions (residues 24-68,92-139,165-187) | 1.26 ± 0.13 Å |
| WHAT IF Z-scores (residues 1-72, 93-155, 165-190)                   |  |
| 1st generation packing quality                                     | 1.19 ± 0.42 |
| 2nd generation packing quality                                     | 3.97 ± 0.87 |
| $\chi^1 \chi^2$ rotamer normality                                  | -6.96 ± 0.27 |
| Backbone conformation                                              | 0.84 ± 0.28 |
Figure 1
Figure 2
Figure 3
Figure 4

IL-6

human TSGSRIDQRYVIRIGSARLKEITCNKSNKETSSKEAENNLHKKMAERDGFQSGFGETGVKIIIGLLEVEVTLEYLCREFEGSE
mouse YSSQQILDHAWWITEYKRKELQIENGSDAENNDALANNKKECTQRNCGQICYNQIDLGISQILYKHNLKDLK

IL-6Rα

human EYRAVQSMVLEQFLQXKAR-NLDAITCPTIVNDRSIT-IKQAAGNCOSQSMTHLIREKKEQQQSDMRAK
mouse DARYLVDSTSGIPNOCEYDLHKIFSCITGAI-KTDKQSENVKQIKKPT-TKLEDFEVKTVSTG

gp130

human RKKFIDQVCEGQGRSTTLPITKAVLFLRFQNSP-ALDFCEGCQYQSSQKQSCQLAQPEGCGFICVSMCNASVGSKFQCKTGQDCG
mouse RRNYHVAICERKVSSTTSPLPTKAVRFAKININNGKPSFFQVCGYQSSLKSCQVHEILGNKTYSECVLWAKNVSQSKKNEASRLK
human IQDPEFAPNITAVARNPRMGSIVUQPSNHSGPFPBEFELYRFAERKTIWLYKOLHHCYTVHHMSLIRIVYQVLRLACAEFGQEP
mouse WYQDPEFAPNITAVARNPRADGRIVUQPSNHSGPFPBEFELYRFAERKTIWLYKOLHHCYTVHHMSLIRIVYQVLRLACAEFGQEP

Site IIa: gp130(D2D3)-IL6
Site IIb: gp130(D2D3)-IL6Rα
Site IIIa: gp130(D1)-IL6
Site IIIb: gp130(D1)-IL6Rα
Figure 5
Figure 6

(A) Wild type ELLDP
(B) E23A mutant ALLDP
(C) L24A mutant EALDP
(D) L25A mutant ELADP
Figure 7
Conservation of functional sites on interleukin-6 and implications for evolution of signalling complex assembly and therapeutic intervention

Vaclav Veverka, Terry Baker, Nicholas T. Redpath, Bruce Carrington, Frederick W. Muskett, Richard J. Taylor, Alastair D. G. Lawson, Alistair J. Henry and Mark D. Carr

J. Biol. Chem. published online October 1, 2012

Access the most updated version of this article at doi: 10.1074/jbc.M112.405597

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2012/10/01/M112.405597.DC1