CHARACTERISATION OF HIGH-AFFINITY BINDING MOTIFS FOR THE DISCOIDIN DOMAIN RECEPTOR DDR2 IN COLLAGEN

Antonios D. Konitsiotis[1,4], Nicolas Raynal[2,4], Dominique Bihan[2], Erhard Hohenester[3], Richard W. Farndale[2,5], and Birgit Leitinger[1,5]

From [1]Division of NHLI, [3]Division of Cell and Molecular Biology, Imperial College London, London SW7 2AZ, United Kingdom [2]Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, United Kingdom

[4,5]These authors contributed equally to this work

Running title: High-affinity DDR2 binding sites in collagen

Address correspondence to: Dr B Leitinger, Molecular Medicine Section, Division of NHLI, Imperial College London, Sir Alexander Fleming Building, London SW7 2AZ, United Kingdom; Tel. +44-20-7594 1591; Fax +44-20-7594 3100; Email: b.leitinger@imperial.ac.uk

The discoidin domain receptors, DDR1 and DDR2, are receptor tyrosine kinases that are activated by native triple-helical collagen. Here we have located three specific DDR2 binding sites by screening the entire triple-helical domain of collagen II, using the Collagen II Toolkit, a set of overlapping triple-helical peptides. The peptide sequence that bound DDR2 with highest affinity interestingly contained the sequence for the high-affinity binding site for von Willebrand factor in collagen III. Focussing on this sequence, we used a set of truncated and alanine-substituted peptides to characterise the sequence GVMGFO (O is hydroxyproline) as the minimal collagen sequence required for DDR2 binding. Based on a recent NMR analysis of the DDR2 collagen-binding domain, we generated a model of the DDR2-collagen interaction that explains why a triple-helical conformation is required for binding. Triple-helical peptides comprising the DDR2 binding motif not only inhibited DDR2 binding to collagen II, but also activated DDR2 transmembrane signalling. Thus, DDR2 activation may be effected by single triple-helices rather than fibrillar collagen.

The mammalian discoidin domain receptors (DDRs), DDR1 and DDR2, are receptor tyrosine kinases (RTKs) that function as collagen receptors (1,2). Several collagen types, in particular fibrillar collagens, bind to and activate the DDRs, with the two receptors displaying different specificities towards certain collagen types (3,4). DDR activation by collagen is strictly dependent on the native, triple-helical conformation of collagen (1,2,5). The DDRs are unique among RTKs in being activated by a component of the extracellular matrix; most RTKs are activated by small diffusible proteins such as growth factors. Like conventional RTKs, the DDRs regulate fundamental cellular processes including cell proliferation, adhesion, and migration, but the DDRs additionally control remodelling of the extracellular matrix (6-9). Both receptors control developmental processes, such as mammary gland development (DDR1) (10) and the growth of long bones (DDR2) (11), and are associated with human diseases, including fibrotic diseases of the liver, kidney and lung, atherosclerosis, osteoarthritis and several types of cancer (reviewed in ref 12).

The homologous DDRs are composed of an N-terminal discoidin homology (DS) domain, followed by a stalk region unique to DDRs (~220 amino acids), a transmembrane domain, a large cytosolic juxtamembrane domain and a C-terminal tyrosine kinase domain (13). DDR activation, manifested by autophosphorylation, is a consequence of collagen binding to a specific site in the DS domain (5,14). Collagen-induced DDR autophosphorylation is unusually slow and sustained (1,2), compared with the rapid response of canonical RTKs to their ligands.

Collagens form a large protein family that is characterised by repeating glycine-X-X' sequences, where X and X' are often proline and 4-hydroxyproline, respectively (15). Three collagen chains form a right-handed triple-helical structure. The fibrillar collagens (types I, II, III, V and XI) are characterised by triple-helical domains of ~1000 amino acids that self-assemble into highly organised fibrils, and play key architectural roles in connective tissues.

The interactions of fibrillar collagens with some of their cellular receptors, in particular integrins, have been well characterised. Collagen-binding integrins recognise discrete amino acid sequences in triple-helical collagen. The use of synthetic triple-helical peptides allowed the identification of an important, high-affinity integrin-binding site, GFOGER (where O is hydroxy-proline) (16,17), which is present in a number of different collagens, including collagens I and II. In contrast to the collagen-binding integrins, for which a receptor-ligand co-crystal has been described as a complex between the alpha-2 I domain and the model collagen peptide GFOGER (18), we know little about the molecular mechanism by which the DDRs recognise collagen. Our previous work showed that DDR2 binds to a specific site within the D2 period of collagen II (3), but the nature of this (or, indeed, any other) DDR2 binding motif is unknown.

In the present study, we have characterised DDR2 binding motifs in collagen, using the collagen II Toolkit, a set of 56 overlapping triple-helical peptides encompassing the entire triple-helical domain of human collagen II. The Toolkit approach allows a comprehensive analysis of binding sites within a fibrillar collagen type. We previously used the collagen III Toolkit to locate novel binding sites for integrins and von Willebrand factor (VWF) in human collagen III (19,20). Screening the collagen II Toolkit for binding to recombinant DDR2 identified three major binding sites. Coincidentally, one of the binding sites is almost identical to the recently identified high-affinity binding site for VWF in collagens II and III (20). Focussing on this peptide sequence, we used a set of truncated and alanine-substituted peptides to characterise the minimal collagen sequence required for DDR2 binding. The specific amino acid motif recognised by DDR2 is distinct from that of integrins and the platelet collagen receptor glycoprotein VI (GpVI) (21). Importantly, triple-helical peptides comprising the DDR2 binding site not only inhibited DDR2 binding to collagen II, but were able to activate DDR2 autophosphorylation in a specific manner.

Experimental Procedures

Peptide synthesis – The sequences of the peptides used in this study are shown in Supplementary Table 1 and in Table 1. Peptides were synthesised by Fmoc (N-(9-fluorenyl)methoxycarbonyl) chemistry as C-terminal amides on TentaGel R RAM resin in an Applied Biosystems Pioneer automated synthesiser and purified as described (19). All peptides were verified by mass spectrometry and shown to adopt triple-helical conformation by polarimetry (Slatter et al, submitted).

Chemicals and reagents – BSA was obtained from Fisher Scientific, Loughborough, UK. Collagen I (acid-soluble from rat tail, C-7661) and collagen IV (acid-soluble from placenta; C-5533) were purchased from Sigma (Poole, UK). Bovine collagen II, ELISA grade, was from Chondrex Inc. (Redmond, WA). The antibodies and their sources were as follows: anti-DDR2, goat anti-DDR2 Ig (AFA2538) from R&D Systems (Abingdon, UK); mouse anti-Myc tag, clone 9E10, from Upstate (Lake Placid, NY)*;* peroxidase-conjugated goat anti-human Fc from Jackson ImmunoResearch Laboratories (West Grove, PA); anti-phosphotyrosine, clone 4G10, from Upstate. Secondary antibodies were as follows: sheep anti-mouse Ig-horseradish peroxidase (Amersham Biosciences UK, Chalfont St Giles, UK); rabbit-anti-goat Ig-horseradish peroxidase (Zymed Laboratories, San Francisco, CA).

Cell culture and cell lines – Human embryonic kidney (HEK) 293 cells (ATCC, Manassas, VA), HEK293-EBNA cells (Invitrogen Ltd, Paisley, UK), and HEK293-T cells (ATCC) were cultured in Dulbecco's modified Eagle medium (DMEM)/F12 nutrient mixture (Invitrogen, UK) with 10% fetal bovine serum.

Production and purification of recombinant proteins – The production and purification of recombinant DDR proteins was as previously described (5). The His-tagged DDR2 extracellular domain protein (His-DDR2) was produced from episomally transfected HEK293-EBNA cells. The Fc-tagged DDR2 DS domain (DS2-Fc) was isolated from episomally transfected HEK293-T cells.

Solid-phase collagen binding assays – The assay procedure has been described previously in detail (5). Briefly, collagens or collagen peptides were diluted in 0.01 M acetic acid to 10

μg/ml and coated onto Immulon 2 HB 96-well plates (Fisher Scientific) overnight at room temperature. This represents a 10-fold excess over saturating concentration of peptide, verified using similar peptides for integrin and platelet binding (D.J. Onley, PhD Thesis, University of Cambridge 2001). Wells were then blocked for 1 h at room temperature with 1 mg/ml BSA in PBS plus 0.05% Tween-20 (PBS-T). Recombinant DDR2 proteins, diluted in incubation buffer (0.5 mg/ml BSA in PBS-T) were added for 3 h at room temperature. Wells were washed 6 times with incubation buffer between all incubation steps. Bound His-DDR2 protein was detected with anti-Myc monoclonal antibody (1:1000 dilution), added for 1 h at room temperature, followed by sheep anti-mouse horseradish peroxidase (1:1000 dilution), added for 1 h at room temperature. Bound DS2-Fc protein was detected with goat anti-human Fc coupled to horseradish peroxidase (1:3333 dilution), added for 1 h at room temperature. A colour reaction was subsequently performed using *o*-phenylenediamine dihydrochloride (Sigma), added for 3-5 min. The reaction was stopped with 3 M H_2SO_4, and plates were read in an ELISA assay reader at 492 nm. For the blocking assays, His-DDR2 was pre-incubated with the indicated concentrations of peptides for 1 h at room temperature prior to adding of the His-DDR2/peptide mixture to collagen II coated plates. The assay was developed as described above.

Collagen-induced DDR autophosphorylation – The assay was performed as previously described in detail (5). Briefly, HEK293 cells in 24-well plates were transfected by calcium phosphate precipitation with DDR2 expression vector (5). 24 h later, the cells were incubated with serum-free medium for 16 h. Cells were then stimulated with collagen I, collagen II (both at 10 μg/ml), or different collagen peptides (at 100 μg/ml) for 90 min, or as indicated, at 37°C. Cells were lysed in 1% Nonidet P-40, 150 mM NaCl, 50 mM Tris, pH 7.4, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 50 μg/ml aprotinin, 1 mM sodium orthovanadate, 5 mM NaF. Aliquots of the lysates were analysed by SDS-PAGE, followed by blotting onto nitrocellulose membranes. The blots were probed with mouse anti-phosphotyrosine mAbs, followed by sheep anti-mouse Ig horseradish peroxidase. Detection was by enhanced chemiluminescence (Amersham Biosciences). To strip the blots, membranes were incubated in Antibody stripping solution (Alpha Diagnostic International, San Antonio, Texas) for 10 min at room temperature. The blots were reprobed with goat anti-DDR2 Abs, followed by rabbit anti-goat Ig-horseradish peroxidase.

RESULTS

The collagen II Toolkit – Sequences of collagen II Toolkit peptides are shown in Supplementary Table 1. The host-guest strategy (22) was applied, as in our previous studies (16,23), where the guest (primary) sequence of interest is placed between $(GPP)_5$ hosts, inert flanking sequences which impart triple-helical conformation on the whole peptide. Each Toolkit peptide contains a guest sequence of 27 amino acids, the C-terminal 9 amino acids of which form the first 9 guest amino acids of the next peptide. Thus, the guest sequence of the Toolkit advances 18 amino acids along the triple-helical domain of collagen II with each successive peptide, and a 9 amino acid overlap is included between adjacent peptides.

Mapping DDR2 binding sites within collagen II – Our previous studies used two recombinant soluble extracellular DDR2 proteins to analyse the binding of DDR2 to immobilised collagen ligands (3-5): one construct comprising the entire DDR2 extracellular domain, N-terminally tagged with a His-tag and a Myc epitope (His-DDR2), the other construct consisting mainly of the DDR2 DS domain, C-terminally tagged with an Fc-tag (DS2-Fc). Both of these constructs exhibited high-affinity binding to the fibrillar collagens I and II, but, like full-length DDR2, did not recognise the basement membrane collagen IV. In this study, we used the same recombinant DDR2 proteins and screened the collagen II Toolkit peptides for DDR2 binding (Figure 1). Both DDR2 proteins showed good binding to only four peptides: Toolkit peptides II-13, II-22, II-23, and II-44. The response was similar to (or greater than) that to full-length immobilised collagen II. However, this might be an overestimation, as the peptides likely immobilise at greater molar concentration on the plates due

to their lower molecular weight compared to full-length collagen II. There was no binding, as expected, to collagen IV or BSA. In addition, the control peptide GPP, [GPC-$(GPP)_{10}$-GPC-NH_2], which consists entirely of the peptide 'host' sequence, was not recognised. Similarly, peptide CRP, [GPC-$(GPO)_{10}$-GPC-NH_2], a ligand for the platelet collagen receptor GpVI (24), did not interact with DDR2. In addition to the four strongly interacting peptides, His-DDR2 showed above background binding to three further peptides, II-17, II-38 and II-39.

Solid phase binding assays showed that the DDR2 proteins had highest affinity to immobilised peptide II-22, amongst the four positive peptides identified in the screen. Half maximal binding occurred at ~10 nM added DDR2 proteins (Figure 2).

Identification of the minimal collagen sequence required for DDR2 binding – As the Toolkit is a set of overlapping peptides, it was likely that DDR2 recognised the overlapping sequence of peptides II-22 and II-23. Inspection of the overlap sequence revealed, interestingly, that this sequence (GQOGVMGFO) is almost identical to the recently identified high-affinity binding site for VWF in collagen III (RGQOGVMGFO) (20). In this previous work, we screened the entire collagen III Toolkit for VWF binding and identified a single peptide, III-23, that bound plasma VWF. We tested this peptide and a set of truncated and alanine-substituted peptides derived from this sequence (see Table 1) for DDR2 binding (Figure 3A and Supplementary Figure 1A). The response of both DDR2 proteins to III-23 was similar to their response to II-22. The results of the truncated peptide series showed that DDR2 bound strongly to the peptide sequence GVMGFO at 10 μg/ml (142 nM) DDR2, but did not bind any peptides in which this hexapeptide sequence was truncated; this hexapeptide will therefore be referred to as the minimal binding sequence. However, at lower DDR2 protein concentrations (2 μg/ml, 28 nM), both GVMGFO and the longer sequence GQOGVMGFO gave sub-optimal responses, indicating that these peptides are recognised with lower affinity. Extending the sequence by a further amino acid triplet to GPRGQOGVMGFO gave optimal responses for both DDR2 concentrations. Dose response curves for both GPRGQOGVMGFO and GPSGPRGQOGVMGFO showed that DDR2 recognised these peptides with similar high affinity that was even somewhat greater than the affinity with which DDR2 recognised Collagen II Toolkit peptide, II-22 (Supplementary Figure 2).

DDR2 binding to intact collagen is strictly dependent on the native triple-helical conformation. It was therefore important to test whether DDR2 would recognise a non-helical collagen peptide. A non-helical version (GAP-GPR-GFO; see Table 1) of the high-affinity DDR2 ligand GPRGQOGVMGFO, was non-helical above 8°C as determined by polarimetry (data not shown), and supported no binding by DDR2 (Figure 3A, Supplementary Figure 1A and Supplementary Figure 2), indicating that peptide recognition by DDR2 in our system mirrors recognition of native collagen by DDR2. Moreover, DDR2 did not recognise GFOGER, the high-affinity integrin binding site identified previously by us (16). Alanine scanning showed that two amino acids, M and F, are crucial for DDR2 recognition; substitution of R and O (underlined in GPRGQOGVMGFO) led to a modest reduction in binding, whereas all other X and X' positions could be substituted with no apparent loss of binding (Figure 3B and Supplementary Figure 1B).

In human collagen I, a heterotrimer of two α1 and one α2 chains, a related sequence occurs in the α1 chain (GARGQAGVMGFO) that differs in two amino acids from the high-affinity DDR2 binding sequence in collagen III (one amino acid difference to the motif GARGQOGVMGFO in collagen II). As expected from the results of the alanine-substituted peptides, a triple-helical peptide comprising this sequence was recognised by DDR2, albeit to a somewhat lesser degree at the lower protein concentrations tested (Figure 3B and Supplementary Figure 1B). The sequence in the α2 chain that aligns with this sequence, GARGEOGNIGFO, was recognised much less well, suggesting that the two α1 chains of collagen are the major contributors to the corresponding DDR2 binding site in collagen I.

High-affinity DDR2 binding peptides inhibit DDR2 binding to collagen II – We tested whether the two high-affinity DDR2 binding peptides, GPRGQOGVMGFO and GPSGPRGQOGVMGFO, could inhibit DDR2

binding to immobilised full-length collagen II. Both peptides efficiently inhibited DDR2 binding with half maximal inhibition at ~50 μg/ml (~10 μM) (Figure 4). Importantly, the non-helical version of GPRGQOGVMGFO did not inhibit DDR2 binding.

Activation of transmembrane signalling by high-affinity DDR2 binding peptides – In order to relate the binding of triple-helical collagen peptides to receptor activation of membrane-bound DDR2, we expressed full-length DDR2 transiently in HEK293 cells and examined whether the peptides could induce DDR2 autophosphorylation (Figure 5). Activation by collagen peptides required 50-100 μg/ml peptides in the medium, compared to 10 μg/ml collagen. The reasons for the lower potency of the peptides are unclear at present. Importantly, however, there was a strict correlation between the ability to induce DDR2 autophosphorylation and high-affinity DDR2 binding. While peptides encompassing the sequence GPRGQOGVMGFO activated the receptor, the two smaller (lower affinity) peptides GVMGFO and GQOGVMGFO did not. Significantly, replacing F by A in the context of otherwise activating sequence GPSGPRGQOGVMGFO abolished DDR2 autophosphorylation. Moreover, the non-helical version of GPRGQOGVMGFO did not activate DDR2. Taken together, these results show that triple-helical peptides comprising the DDR2 binding site activate DDR2 autophosphorylation in a specific manner, demonstrating that the identified triple-helical motif is a genuine DDR agonist leading to transmembrane signalling.

Collagen-induced DDR autophosphorylation occurs with unusually slow and sustained kinetics (1,2). Peptide GPRGQOGVMGFO induced DDR2 autophosphorylation with the same slow kinetics as collagen I, with maximal DDR2 phosphorylation detectable at ~60 min of incubation (Figure 6). This result indicates that slow cellular processes are responsible for the delay in receptor activation, rather than accessibility of ligand binding sites within collagen.

DISCUSSION

A major unresolved question is whether the DDRs recognise specific sequence motifs in collagen (13). Here we define, for the first time, specific binding motifs for DDR2 in collagen. Our results show that, like collagen-binding integrins, DDR2 recognises discrete amino acid motifs in triple-helical collagen. DDR2 binding motifs are distinct from those defined for integrins (summarised by ref (19) and the platelet collagen receptor GpVI (21,24). Screening the Collagen II Toolkit for DDR2 binding identified only four peptides that bound DDR2 strongly. The three binding sites identified in collagen II each contain a GFO triplet, with the F critical and the O important for DDR2 binding to the motif contained in the overlap of Toolkit peptides II-22 and II-23 (henceforth referred to as the central DDR2 binding motif). The presence of GFO is not sufficient for DDR2 binding, however: there are eight GFO motifs in the human collagen II sequence, five of which do not bind DDR2. Importantly, one of these GFO triplets is part of the integrin-binding motif, GFOGER, which DDR2 did not recognise, neither as an isolated motif nor in peptide II-28, the Toolkit peptide that contains this motif.

Our previous work identified a major DDR2 binding site in the D2 period of collagen II (amino acids 235-468 of the triple-helical COL 1 domain, corresponding to residues 366-599 of the human collagen II precursor, Swiss-Prot entry P02458) (3). The central DDR2 binding motif characterised here in detail corresponds to amino acids 391-405 of human collagen II, corroborating our earlier findings. The other two DDR2 binding sites found here, contained in Toolkit peptides II-13 and II-44, correspond to amino acids 217-243 and 775-801, respectively. Deletion of the D1 period of collagen II (amino acids 1-234) partially diminished DDR2 binding and inhibited DDR2 autophosphorylation (3), corresponding well with our present findings for peptide II-13. However, our previous work did not identify DDR2 binding sites in the D4 period of collagen II (amino acids 703-936).

Of the three DDR2 binding sites in collagen II, only the central motif is conserved in collagen III (Figure 7). However, sequences corresponding to all three binding sites are conserved in the α1 chain of collagen I. DDR2 bound a peptide from the α1 chain of collagen I corresponding to the central binding motif (Figure 3 and Supplementary Figure 1). This

peptide was also able to induce DDR2 autophosphorylation (data not shown). Together, these findings imply that DDR2 recognises mainly the α1 chains in collagen I.

The central binding motif for DDR2 overlaps with that for VWF in collagen III (RGQOGVMGFO). However, the amino acid requirements for DDR2 recognition differ substantially from those previously identified for VWF recognition (20) (for summary see Figure 7). While F is crucial for both DDR2 and VWF, the latter has an absolute requirement for R (only partially required for DDR2), the first O and V (neither required for DDR2), while DDR2 requires M, which is not required for VWF binding. This overlap of binding sites within collagens suggests that ligands may compete for collagen exposed during trauma or disease. It is plausible that plasma-borne VWF may inhibit the DDR2-collagen interaction, thus restricting the activation of DDR2-expressing cells to sites remote from the circulation.

The minimum collagen sequence that bound DDR2 comprised six amino acids, GVMGFO, like the canonical GXX'GEX" motif recognised by integrin I domains (16,19,25-27). F and M were found to be critical for the interaction of DDR2 with the GVMGFO motif. While this manuscript was in preparation, the solution structure of the DDR2 DS domain was reported and the collagen-binding site mapped by transferred cross-saturation and mutagenesis experiments (14). The binding site is created by surface-exposed loops that form a trench suitable to accommodate the collagen triple helix (Figure 8). The length of this trench spans two to three triplets in collagen, in good agreement with our results. The residues making up the trench include charged amino acids, as well as a number of polar and apolar residues. The most striking feature is a solvent-exposed aromatic residue, W52, whose mutation to alanine essentially abrogated collagen II binding (14).

To gain insight into the mode of collagen binding to the DDR2 DS domain, we created a molecular model of the DDR2-binding sequence GPRGQOGVMGFOGPK based on a $(GPO)_n$ backbone of 10/3 helical symmetry (Figure 8). From this model, it is immediately evident why a single polypeptide chain does not bind to DDR2: the critical M and F side chains project from opposite sides of the chain and cannot interact simultaneously with the trench on the DS domain. In the triple helix, however, M and F side chains from different chains come into close proximity, and we propose that one of the apolar patches thus formed is accommodated by the hydrophobic region of the DS domain trench, in particular by the critical side chain of W52 and the C73-C177 disulphide bond at the floor of the trench. Because of the stagger between individual chains of a collagen triple helix, there exist two different arrangements of M and F side chains: one type occurs only once per triple helix and involves F of the leading strand and M of the trailing strand (facing left in Figure 8); the second type occurs twice and involves M from the leading or middle strand and F from the middle or trailing strand, respectively (facing right and to the back in Figure 8). We used manual and computational docking to see whether one of these types led to a preferred mode of binding (data not shown). However, both types could be accommodated equally well, with either M or F (or possibly O) stacking against W52. Whether this uncertainty is due to the limitations of the docking routines (we did not attempt to model conformational changes in the DS domain) or an inherent feature of the DDR2-collagen interaction remains to be seen. A unique binding mode may be imposed by interactions outside the core GVMGFO motif in collagen. Indeed, our data show that the arginine residue upstream of the hydrophobic motif contributes to binding and DDR2 activation. In a subset of our docking models, this arginine can be oriented to make an electrostatic interaction with E113, which has been implicated in collagen binding (14). Further structural studies will be required to define the atomic details of the DDR2-collagen interaction.

DDR2 has the same overall architecture as the homologous DDR1. The DS domains of the DDRs share an overall sequence identity of 59%, but the surface-exposed loops comprising the collagen binding site of DDR2 are strikingly conserved (5), with 11 of the 13 amino acids identified by NMR identical between DDR1 and DDR2 (14). Despite this sequence similarity, the two receptors bind collagen with different specificities: DDR1 binds to the basement membrane collagen IV, but DDR2 does not recognise this ligand (1,2,5); DDR2 binds collagens II and X with much higher affinity

than DDR1 (3,4); both DDRs bind collagen I with similar high-affinity (5). Regarding the latter activity, our preliminary experiments have shown that DDR1 binding to collagen I can be competed with excess DDR2 (data not shown), indicating that the DDRs might have overlapping binding sites on collagen I. Based on the sequence conservation of the DDR2 loops that are critical for collagen I binding, we suggest that the DDRs recognise similar collagen sequences, but that they bind these motifs in slightly different ways.

We found that triple-helical DDR2-binding peptides not only blocked DDR2 binding to collagen II, but were specific agonists for the receptor. As these activating peptides are single triple helices in solution, it follows that the higher-order organisation of collagen into fibrils or fibres is not required for DDR2 activation. These findings may offer a possible solution to the conundrum of DDR regulation, given the abundance of collagen fibrils and fibres in tissues in which the DDRs are expressed: the receptors may be sensitive only to single triple helices, and not to fibrils or fibres. Single triple helices may be more exposed in disease states or during tissue remodelling, and the DDRs have indeed been found to be involved in tissue remodelling, such as wound healing (11), cancer invasion (28,29) or fibrosis (7,30,31). However, it remains to be established whether the DDRs can be activated by intact collagen fibrils or fibers. In contrast to the DDRs, the activation of other collagen receptors is known to require higher-order collagen structure. For example, GpVI, which recognises the peptide CRP [GPC-$(GPO)_{10}$-GPC-NH_2], can only be activated after peptide cross-linking to introduce polymeric structures (24). This implies a different mechanism of signalling in DDR2 than in GpVI: the latter appears to require a multivalent ligand that can induce receptor clustering to signal, although the receptor will recognise and adhere to monomeric ligands (21). The same property has been reported for another immunoglobulin superfamily collagen receptor, LAIR-1 (32).

The mechanism of DDR activation also differs from that of canonical RTKs, which are thought to be activated by ligand-induced dimerisation (33). In contrast to these receptors, the DDRs are predimerised at the cell membrane in the absence of ligand, and collagen binding is predicted to lead to a conformational change within the DDR dimer (34). The ~1000-fold difference in collagen affinity between monomeric (14) and dimerised DS domain (5) argues against two independent collagen binding sites, and the signalling complex may consist of one triple helix per DDR dimer. The collagen binding trench on the DS domain is rather shallow and would leave part of the triple helix available for recognition by a second DS domain. Since the DDR dimer is likely to have two-fold rotational symmetry, and a collagen triple helix is lacking such symmetry, the DS domain would have to be able to accommodate the triple helix in both directions. In this regard, we are intrigued by the absence of features in the DDR2-binding collagen motif imparting a strong directionality (see above). Perhaps, what is being recognised by the DDRs is the rare concentration of large hydrophobic side chains, which is found on opposite faces of the GVMGFO triple helix (Figure 8).

In summary, our comprehensive analysis of DDR2 binding sites in human collagen II has identified three high-affinity binding sites, which are distinct from integrin and GpVI interaction sites. One of these binding sites is conserved between the α1 chains of collagens I, II and III, and overlaps, but is not identical, with the previously identified binding motif for VWF in collagen III. High-affinity peptides comprising the DDR2 binding site activated the receptor in a conformation-dependent manner, suggesting new mechanisms of receptor signalling and offering a new perspective on DDR activation *in vivo*. The specific DDR2-binding collagen sequence described here will be an invaluable tool for further studies into DDR function.

REFERENCES

1. Vogel, W., Gish, G. D., Alves, F., and Pawson, T. (1997) *Mol Cell* **1**, 13-23

2. Shrivastava, A., Radziejewski, C., Campbell, E., Kovac, L., McGlynn, M., Ryan, T. E., Davis, S., Goldfarb, M. P., Glass, D. J., Lemke, G., and Yancopoulos, G. D. (1997) *Mol Cell* **1**, 25-34
3. Leitinger, B., Steplewski, A., and Fertala, A. (2004) *J Mol Biol* **344**, 993-1003
4. Leitinger, B., and Kwan, A. P. (2006) *Matrix Biol* **25**, 355-364
5. Leitinger, B. (2003) *J Biol Chem* **278**, 16761-16769
6. Hou, G., Vogel, W., and Bendeck, M. P. (2001) *J Clin Invest* **107**, 727-735.
7. Olaso, E., Ikeda, K., Eng, F. J., Xu, L., Wang, L. H., Lin, H. C., and Friedman, S. L. (2001) *J Clin Invest* **108**, 1369-1378
8. Olaso, E., Labrador, J. P., Wang, L., Ikeda, K., Eng, F. J., Klein, R., Lovett, D. H., Lin, H. C., and Friedman, S. L. (2002) *J Biol Chem* **277**, 3606-3613
9. Ferri, N., Carragher, N. O., and Raines, E. W. (2004) *Am J Pathol* **164**, 1575-1585
10. Vogel, W. F., Aszodi, A., Alves, F., and Pawson, T. (2001) *Mol Cell Biol* **21**, 2906-2917.
11. Labrador, J. P., Azcoitia, V., Tuckermann, J., Lin, C., Olaso, E., Manes, S., Bruckner, K., Goergen, J. L., Lemke, G., Yancopoulos, G., Angel, P., Martinez, A. C., and Klein, R. (2001) *EMBO Rep* **2**, 446-452.
12. Vogel, W. F., Abdulhussein, R., and Ford, C. E. (2006) *Cell Signal* **18**, 1108-1116
13. Leitinger, B., and Hohenester, E. (2007) *Matrix Biol* **26**, 146-155
14. Ichikawa, O., Osawa, M., Nishida, N., Goshima, N., Nomura, N., and Shimada, I. (2007) *Embo J* **26**, 4168-4176
15. Kadler, K. E., Baldock, C., Bella, J., and Boot-Handford, R. P. (2007) *J Cell Sci* **120**, 1955-1958
16. Knight, C. G., Morton, L. F., Peachey, A. R., Tuckwell, D. S., Farndale, R. W., and Barnes, M. J. (2000) *J Biol Chem* **275**, 35-40
17. Zhang, W. M., Kapyla, J., Puranen, J. S., Knight, C. G., Tiger, C. F., Pentikainen, O. T., Johnson, M. S., Farndale, R. W., Heino, J., and Gullberg, D. (2003) *J Biol Chem* **278**, 7270-7277
18. Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J., and Liddington, R. C. (2000) *Cell* **101**, 47-56
19. Raynal, N., Hamaia, S. W., Siljander, P. R., Maddox, B., Peachey, A. R., Fernandez, R., Foley, L. J., Slatter, D. A., Jarvis, G. E., and Farndale, R. W. (2006) *J Biol Chem* **281**, 3821-3831
20. Lisman, T., Raynal, N., Groeneveld, D., Maddox, B., Peachey, A. R., Huizinga, E. G., de Groot, P. G., and Farndale, R. W. (2006) *Blood* **108**, 3753-3756
21. Smethurst, P. A., Onley, D. J., Jarvis, G. E., O'Connor, M. N., Knight, C. G., Herr, A. B., Ouwehand, W. H., and Farndale, R. W. (2007) *J Biol Chem* **282**, 1296-1304
22. Shah, N. K., Ramshaw, J. A., Kirkpatrick, A., Shah, C., and Brodsky, B. (1996) *Biochemistry* **35**, 10262-10268
23. Knight, C. G., Morton, L. F., Onley, D. J., Peachey, A. R., Messent, A. J., Smethurst, P. A., Tuckwell, D. S., Farndale, R. W., and Barnes, M. J. (1998) *J Biol Chem* **273**, 33287-33294
24. Morton, L. F., Hargreaves, P. G., Farndale, R. W., Young, R. D., and Barnes, M. J. (1995) *Biochem J* **306**, 337-344
25. Xu, Y., Gurusiddappa, S., Rich, R. L., Owens, R. T., Keene, D. R., Mayne, R., Hook, A., and Hook, M. (2000) *J Biol Chem* **275**, 38981-38989
26. Siljander, P. R., Hamaia, S., Peachey, A. R., Slatter, D. A., Smethurst, P. A., Ouwehand, W. H., Knight, C. G., and Farndale, R. W. (2004) *J Biol Chem* **279**, 47763-47772
27. Kim, J. K., Xu, Y., Xu, X., Keene, D. R., Gurusiddappa, S., Liang, X., Wary, K. K., and Hook, M. (2005) *J Biol Chem* **280**, 32512-32520
28. Ram, R., Lorente, G., Nikolich, K., Urfer, R., Foehr, E., and Nagavarapu, U. (2006) *J Neurooncol* **76**, 239-248
29. Yoshida, D., and Teramoto, A. (2007) *J Neurooncol* **82**, 29-40
30. Matsuyama, W., Watanabe, M., Shirahama, Y., Oonakahara, K., Higashimoto, I., Yoshimura, T., Osame, M., and Arimura, K. (2005) *J Immunol* **174**, 6490-6498
31. Flamant, M., Placier, S., Rodenas, A., Curat, C. A., Vogel, W. F., Chatziantoniou, C., and Dussaule, J. C. (2006) *J Am Soc Nephrol* **17**, 3374-3381

32. Lebbink, R. J., de Ruiter, T., Kaptijn, G. J., Bihan, D. G., Jansen, C. A., Lenting, P. J., and Meyaard, L. (2007) *Int Immunol* **19**, 1011-1019
33. Schlessinger, J. (2000) *Cell* **103**, 211-225
34. Noordeen, N. A., Carafoli, F., Hohenester, E., Horton, M. A., and Leitinger, B. (2006) *J Biol Chem* **281**, 22744-22751

FOOTNOTES

Acknowledgements – The Wellcome Trust and MRC supported the synthesis of peptides. EH is a Wellcome Trust Senior Fellow.

The abbreviations used are: DDR, discoidin domain receptor; DS, discoidin homology; GpVI, glycoprotein VI; RTK, receptor tyrosine kinase; VWF, von Willebrand factor

FIGURE LEGENDS

Figure 1. Identification of DDR2 binding sites on collagen II. Binding of recombinant DDR2 extracellular domain proteins to immobilised collagen II Toolkit peptides in a solid phase binding assay. Recombinant DDR2 proteins were added for three hours at room temperature to 96 wells coated with collagen or peptides at 10 μg/ml. Bound proteins were detected with antibodies and measured as absorbance at 492 nm. (A) Binding of DDR2 extracellular domain, His-DDR2, added at 10 μg/ml (142 nM). (B) Binding of DDR2 DS domain protein, DS2-Fc, added at 10 μg/ml (159 nM). Shown are the mean +/- SD of three independent experiments, each performed in triplicates. CII: bovine collagen II; CIV: human placental collagen IV, GPP and CRP: peptides as shown in Supplementary Table 1.

Figure 2. DDR2 recognises collagen II Toolkit peptide, II-22, with high affinity. Solid phase binding assay with recombinant DDR2 proteins, added for three hours at room temperature to 96 wells coated with triple-helical collagen peptides or collagen at 10 μg/ml. (A) His-DDR2, (B) DS2-Fc binding to different immobilised peptides: Collagen II Toolkit peptide, II-22 (▲); bovine collagen II (△); collagen II Toolkit peptide, II-44 (●); collagen II Toolkit peptide, II-23 (○); collagen II Toolkit peptide, II-13 (◆). Shown is a representative of five independent experiments, each performed in duplicates.

Figure 3. Characterisation of collagen sequence required for high-affinity DDR2 binding. Recombinant His-DDR2 protein was added for three hours at room temperature to 96 wells coated with collagen-derived peptides or collagen at 10 μg/ml. (A) Binding of His-DDR2 to collagen III Toolkit peptide, III-23, and truncated peptides derived from this sequence (Table 1). (B) Binding of His-DDR2 to alanine-substituted peptides derived from collagen III Toolkit peptide, III-23 (Table 1). *CII*, full length bovine collagen II. Shown are the mean +/- SD of three independent experiments, each performed in triplicates. Black bars, His-DDR2 added at 10 μg/ml (142 nM); grey bars, His-DDR2 added at 2 μg/ml (28 nM).

Figure 4. High affinity DDR2 binding peptides block DDR2 binding to collagen II. His-DDR2 was added at 43 nM for three hours to 96 wells coated with 10 μg/ml bovine collagen II in the presence or absence of different concentrations of non-helical or triple-helical collagen peptides: non-helical peptide GPC-$(GAP)_5$-GPRGQOGVMGFO-$(GAP)_5$GPC-NH_2 (●); triple-helical peptide GPC-$(GPP)_5$-GPRGQOGVMGFO-$(GPP)_5$-GPC-NH_2 (○); triple-helical peptide GPC-$(GPP)_5$-GPSGPRGQOGVMGFO-$(GPP)_5$-GPC-NH_2 (▲). Shown is the mean +/- SD of three independent experiments, each performed in triplicates, normalised to His-DDR2 binding in the absence of peptides, which was set to an absorbance of 1.0.

Figure 5. DDR2 binding peptides mediate autophosphorylation of cell surface DDR2. Full length DDR2 was transiently expressed in HEK293 cells. After stimulation for 90 min with collagen at 10 μg/ml or collagen peptides at 100 μg/ml, cell lysates were analysed by SDS-PAGE and Western blotting. *CI,* rat tail collagen I; *CII*, bovine collagen II; peptide names refer to Table 1. The blot was probed with anti-phosphotyrosine mAb 4G10 (upper panel), followed by stripping and reprobing with anti-DDR2 (lower panel). The position of molecular weight markers (in kDa) are indicated. The experiment was carried out four times with very similar results.

Figure 6. DDR2 activating peptide induces receptor phosphorylation with the same slow kinetics as full-length collagen. Full length DDR2 was transiently expressed in HEK293 cells. Cells were stimulated with either rat tail collagen I or triple-helical peptide GPC-$(GPP)_5$-GPRGQOGVMGFO-$(GPP)_5$-GPC-NH_2 for the indicated amounts of time (minutes). Cell lysates were analysed by SDS-PAGE and Western blotting. The blot was probed with anti-phosphotyrosine mAb 4G10 (upper panel), followed by stripping and reprobing with anti-DDR2 (lower panel). The position of molecular weight markers (in kDa) are indicated. The experiment was carried out three times with very similar results.

Figure 7. Alignment of human collagen sequences corresponding to collagen II Toolkit peptides II-13, II-22/23 and II-44. *CI; CII; CIII*: α1 chains of collagens I, II and III. Amino acids important in DDR2 (*) and VWF (º) recognition sites in collagen III are marked.

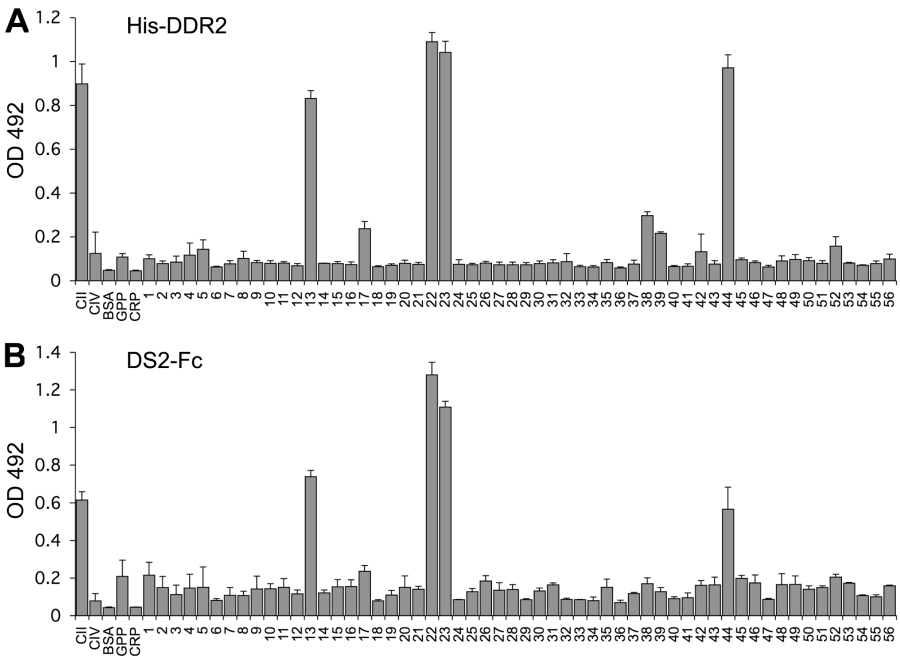
Figure 8. Molecular structures of the DDR2 DS domain and its binding site on collagen II. The structure of the DS domain (14) is shown as a surface representation, with residues implicated in collagen binding coloured and labelled (red, acidic residues; blue, basic residues; cyan, polar residues; green, hydrophobic residues). The collagen triple helix is shown with its three constituent chains distinguished by colouring (yellow, leading chain; orange, middle chain; red, trailing chain; chains are defined as in (18). Side chains important for DDR2 binding are shown as stick models and are labelled. The side chain conformation of the phenylalanine (F) is restricted to a single rotamer by the collagen structure; other side chains have been modelled to be energetically favourable. The hydrophobic ring around the collagen helix is indicated by square brackets and proposed to bind in the hydrophobic depression lined by DS residues W52, C73, N175 and C177 (see text).

Table 1
Guest sequences of peptides used to characterise DDR2-collagen interaction[1]

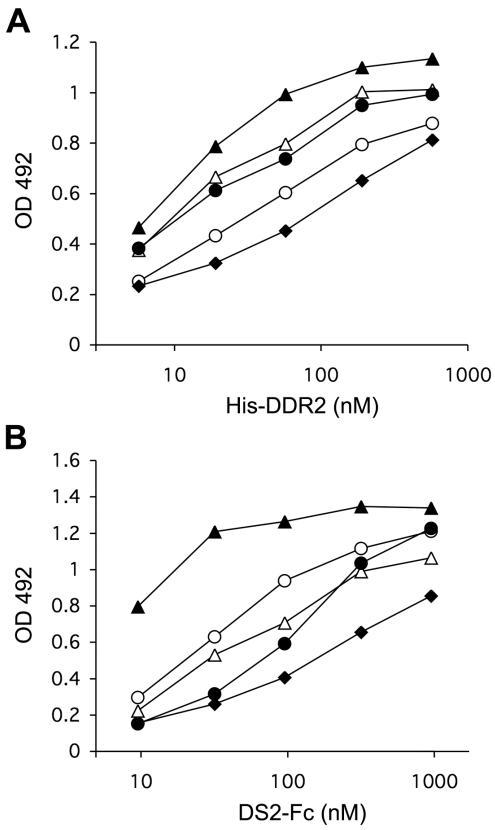
Peptide Name	Sequence	DDR2 binding
II-22	GAOGEDGROGPOGPQGARGQOGVMGFO	+
II-23	GQOGVMGFOGPKGANGEOGKAGEKGLO	+
III-23	GPOGPSGPRGQOGVMGFOGPKGNDGAO	+
GPS-GFO	GPSGPRGQOGVMGFO	+
GPS-GVM	GPSGPRGQOGVM	–
GPS-GQO	GPSGPRGQO	–
GFOGPK	GFOGPK	–
GPR-GPK	GPRGQOGVMGFOGPK	+
GPR-GFO	GPRGQOGVMGFO	+
GQO-GFO	GQOGVMGFO	+/–
GVMGFO	GVMGFO	+/–
GAP-GPR-GFO[2]	$(GAP)_5$-GPRGQOGVMGFO-$(GAP)_5$	–
GFOGER	GFOGER	–
GPS-GFO-5A	GPSG**A**RGQOGVMGFO	+
GPS-GFO-6A	GPSGP**A**GQOGVMGFO	+/–
GPS-GFO-8	GPSGPRG**A**OGVMGFO	+
GPS-GFO-9A	GPSGPRGQ**A**GVMGFO	+
GPS-GFO-11A	GPSGPRGQOG**A**MGFO	+
GPS-GFO-12A	GPSGPRGQOGV**A**GFO	–
GPS-GFO-14A	GPSGPRGQOGVMG**A**O	–
GPS-GFO-15A	GPSGPRGQOGVMGF**A**	+/–
CI-alpha1	GARGQAGVMGFO	+
CI-alpha2	GARGEOGNIGFO	–

[1] *DDR2 binding*, denotes binding to His-DDR2 or DS2-Fc. A plus sign (+) indicates a response similar to peptides II-22, III-23 or GPS-GFO (control); a minus sign (-) indicates less than 20% of control; +/- indicates 20-80% of control.
[2] GAP-GPRGFO: non-helical peptide flanked by $(GAP)_5$ instead of $(GPP)_5$ flanking all other peptides.

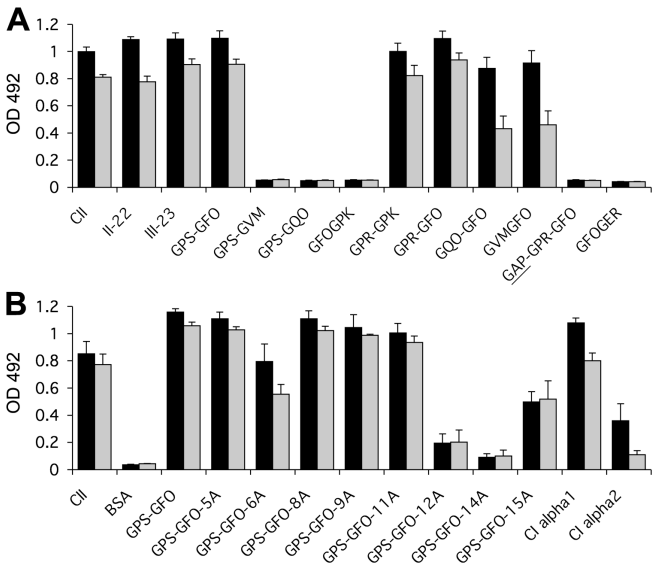
Konitsiotis et al, Figure 1



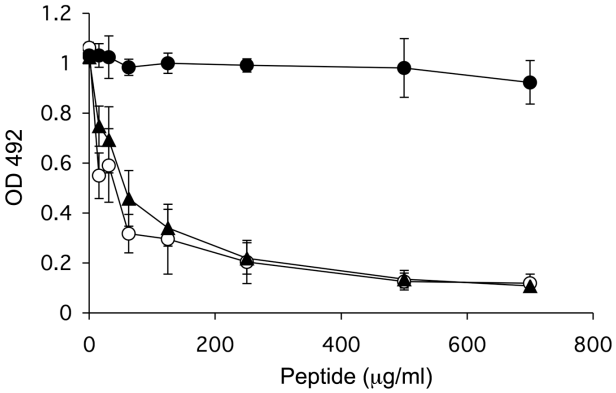
Konitsiotis et al, Figure 2



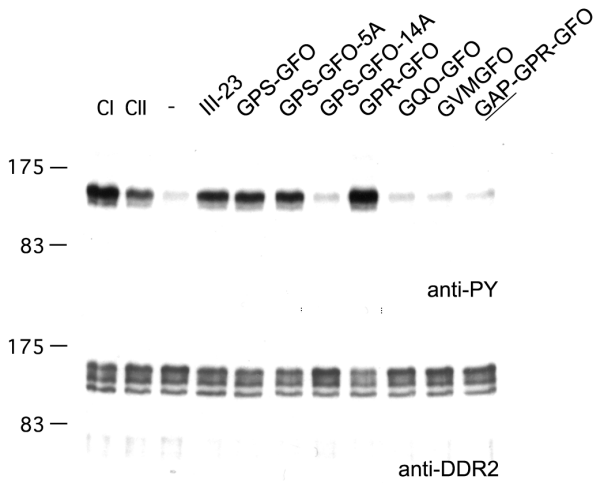
Konitsiotis et al, Figure 3



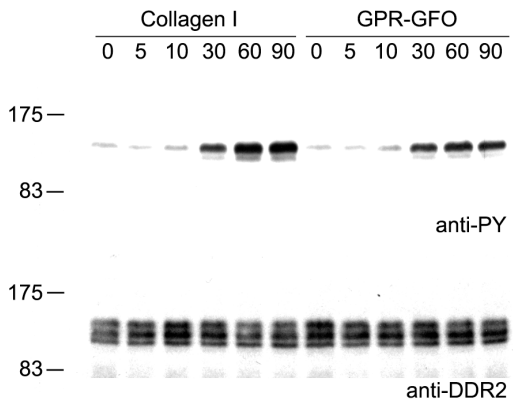
Konitsiotis et al, Figure 4



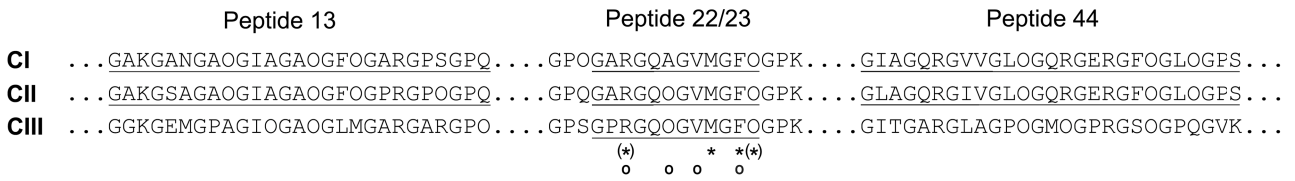
Konitsiotis et al, Figure 5



Konitsiotis et al, Figure 6



Konitsiotis et al, Figure 7



Konitsiotis et al, Figure 8