The Structure of Integrin α1I Domain in Complex with a Collagen-mimetic Peptide*

Received for publication, April 23, 2013, and in revised form, October 4, 2013 Published, JBC Papers in Press, November 1, 2013, DOI 10.1074/jbc.M113.480251

Yanni K.-Y. Chin‡, Stephen J. Headey‡, Biswaranjani Mohanty‡§, Rahul Patil‡, Paul A. McEwan‡, James D. Swarbrick‡, Terrence D. Mulhern¶, Jonas Emsley‡, Jamie S. Simpson‡, and Martin J. Scanlon‡‡

From‡Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences and §Australian Research Council Centre of Excellence for Coherent X-ray Science, Monash University, Parkville, Victoria 3052, Australia, ¶School of Pharmacy, Centre for Biomolecular Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom, and ‡Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Victoria 3010, Australia

Background: Collagen-binding integrins bind differentially to different types of collagen.

Results: The solution structure of integrin α1I domain in complex with a collagen-mimetic peptide was determined.

Conclusion: Integrin α1I domain binds collagen in a distinct orientation compared with α2I, but the signal transduction mechanisms appear to be conserved.

Significance: Understanding the collagen binding specificity of integrins might enable their selective modulation in disease.

We have determined the structure of the human integrin α1I domain bound to a triple-helical collagen peptide. The structure of the α1I-peptide complex was investigated using data from NMR, small angle x-ray scattering, and size exclusion chromatography that were used to generate and validate a model of the complex using the data-driven docking program, HADDOCK (High Ambiguity Driven Biomolecular Docking). The structure revealed that the α1I domain undergoes a major conformational change upon binding of the collagen peptide. This involves a large movement in the C-terminal helix of the αI domain that has been suggested to be the mechanism by which signals are propagated in the intact integrin receptor. The structure suggests a basis for the different binding selectivity observed for the α1I and α2I domains. Mutational data identify residues that contribute to the conformational change observed. Furthermore, small angle x-ray scattering data suggest that at low collagen peptide concentrations the complex exists in equilibrium between a 1:1 and 2:1 α1I-peptide complex.

Integrins comprise a family of non-covalently associated heterodimeric cell surface receptors containing an α subunit and a β subunit. They mediate interactions between individual cells as well as interactions between cells and the extracellular matrix (ECM). These receptors act as mediators that transmit bidirectional signals across the cell membrane. In outside-in signaling, signals are triggered by the binding of ECM ligands to the extracellular domain of the integrin and are propagated through the receptor to reach the intracellular domain and generate cellular responses. In contrast, inside-out signaling results from the binding of cytosolic molecules to the cytoplasmic domain of the integrin receptor, which can regulate the binding affinity of the receptors for ligands in the ECM (1).

Four of the 24 integrins are characterized by their ability to bind to collagen and are categorized as the collagen-binding integrins. All four collagen-binding integrins contain a β1 subunit that is in complex with an α1, α2, α10, or α11 subunit to form the α1β1, α2β1, α10β1, and α11β1 receptors. These four collagen-binding integrins vary in terms of their tissue distribution and cellular signaling pathways, but they share a common role of mechanistically supporting cell adhesion to collagen in the ECM and maintaining tissue integrity (2). Of these, α1β1 and α2β1 are the best characterized and have been reported to play important roles in cellular processes such as angiogenesis, the regulation of collagen expression by fibroblasts, T-cell activation, and collagen-induced platelet aggregation (2–4). They have also been implicated in disease processes including inflammation as well as metastasis in certain cancers, making them attractive therapeutic targets (4–7).

All four collagen-binding integrins contain an inserted domain called the αI domain (αI), comprising ~200 residues, which is located in the extracellular region of the α subunit. The αI domains contain the major binding site for ECM ligands of these receptors, and recombinant isolated αI domains recapitulate many of the ligand binding properties of the intact integrins (8, 9). This makes the αI domains suitable models for the study of specific interactions between integrins and ECM ligands. All αI domains assume a Rossmann fold, which is composed of five parallel β-strands and one antiparallel β-strand.

---

*This work was supported in part by a grant from the Monash-Nottingham Research Fund.

Chemical shifts of all assigned resonances of the GLOGEN-bound α1I were deposited in BioMagResBank under accession number 18942. The atomic coordinates and structure factors (code 2M32) have been deposited in the Protein Data Bank (http://wwpdb.org/).

1 To whom correspondence should be addressed: Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, 381 Royal Parade, Parkville, Victoria 3052, Australia; Tel.: 61-3-99039540; Fax: 61-3-99039582; E-mail: martin.scanlon@monash.edu.

2 The abbreviations used are: ECM, extracellular matrix; HADDOCK, High Ambiguity Driven Biomolecular Docking; αI, αI domain; α2I, αI domain; MIDAS, metal ion-dependent adhesion site; O or Hyp, hydroxyproline; HSQC, heteronuclear single quantum correlation; SOFAST HMQC, selective optimized flip angle short transient heteronuclear multiple quantum coherence; CSP, chemical shift perturbation; TROSY, transverse relaxation-optimized spectroscopy; AIR, ambiguous interaction restraint; SEC, size exclusion chromatography; SAXS, small angle x-ray scattering; WAXS, wide angle x-ray scattering.
surrounded by six α-helices. The αI domain coordinates a divalent metal ion at a site termed the metal ion-dependent adhesion site (MIDAS), which is part of the collagen-binding site. The metal-coordinating residues of the MIDAS are located in three loops on one surface of the αI domain (10, 11).

The four collagen-binding integrins bind differentially to different types of collagen. For example, α1β1 has a higher binding affinity for collagen type IV than for collagen type I, but α2β1 preferentially binds collagen type I over collagen type IV (12–14). Synthetic peptides that mimic specific collagen sequences and contain repeats of the amino acid motif GXX have been used widely in the study of integrin-collagen interactions. These peptides spontaneously self-assemble to adopt the triple-helical conformation that is found in native collagen. As these peptides form homotrimers, they do not recapitulate the heterotrimeric structure observed in some collagens. Nonetheless, they have proved to be extremely valuable tools to probe the binding specificity and structure of collagen-binding proteins (15–18). Using peptides such as these, several specific recognition motifs have been identified for α1β1 and α2β1 integrins (18–21). The recognition motifs are generally six-residue sequences flanked on either side by several repeats of GPO (glycine-proline-hydroxyproline). They include GFOGER (15), which is found in collagen type I, as well as GLOGEN and GROGER from collagen type III (18, 21).

A crystal structure has been reported previously for the α2I domain (α2I) in complex with a triple-helical peptide containing the GFOGER motif (22). Comparison of the liganded and unliganded α2I structures revealed that significant conformational changes occur on binding (22). This structure provided the first insights into the activation mechanism of collagen-binding αI domains. The major structural changes observed on ligand binding include 1) a rearrangement at the MIDAS that was caused by a change in the metal ion coordination involving a glutamate residue from the collagen peptide (GFOGER); 2) a 10-Å downward shift of the C-terminal helix (helix 7), which was proposed to form an interaction with the β1 subunit of the receptor and propagate the signal toward the cytoplasmic domain (22–26); and 3) an uncoiling of a one-turn-and-a-half helix (C helix), which is located above the MIDAS in the unliganded state. The uncoiling was proposed to open up the MIDAS to allow the collagen-mimetic peptide to bind. Ligand binding to α2I was accompanied by the breakage of a salt bridge between Arg287 and Glu317, which is located in the C helix, and residue Glu318 from helix 7. Loss of this salt bridge was proposed to be partly responsible for the uncoiling of the C helix that accompanies peptide binding. A single point mutation at Glu318 (E318W) showed increased binding affinity for collagen, suggesting the involvement of this conformational change in α2I activation (27, 28).

Because of their sequence and structural homology and their shared ability to bind to collagen, α1I and α2I domains have been proposed to use a similar activation mechanism to transmit signals. However, there have been no structures of any α1I-peptide complexes to confirm this hypothesis. The structure of the unliganded form of α1I reveals a salt bridge between residues Arg287 and Glu317, and the structure of α1I containing the mutation E317A was reported in 2011 (29). The E317A mutant showed enhanced binding affinity for collagen as had previously been observed following mutation of the similar residue in α2I (E318W), although the structure was different from that of α2I in complex with GFOGER. The C helix in α1I E317A uncoiled as observed in the α2I-GFOGER complex, but the displacement of helix 7 did not occur, and the metal ion at the MIDAS of α1I E317A was pentacoordinated, whereas the MIDAS in α2I-GFOGER was hexacoordinated. More recently, analysis of the solution structure of E317A using NMR (30) showed evidence of significant conformational change in the mutant. In addition, the E317A mutation led to increased binding to collagen and increased activation of the intact α1β1 receptor. In contrast, the mutation R287A produced less pronounced effects on collagen binding and receptor activation, whereas a charge-reversed mutant, R287E/E317R, displayed a phenotype more similar to E317A. This led the authors to propose that 1) the role of Glu317 in stabilizing the low affinity conformation of α1I is largely independent of the salt bridge with Arg287 and may result from interactions with helix dipoles in the structure and 2) mutation of Glu317 may be sufficient to result in displacement of helix 7 in contrast to the crystal structure (29). However, the NMR data demonstrated that the protein containing the E317A mutation was conformationally flexible, and it is possible that the crystal structure represents one state of E317A.

Here, we report the solution structure of human α1I domain in complex with a collagen-mimetic peptide. The triple-helical peptide contained a GLOGEN recognition sequence, which is a high affinity ligand of α1I (18). The binding orientation of the peptide is different from that observed for GFOGER binding to α2I, and a close comparison of the interfaces reveals differences that may account for their collagen binding preferences. Despite the difference in ligand binding orientations, the signaling mechanisms appear to be conserved. GLOGEN binding to α1I resulted in significant conformation changes including the uncoiling of the C helix, displacement of helix 7, and breakage of the Arg287-Glu317 salt bridge, consistent with what was observed in α2I domain upon binding. The structure of the complex revealed the formation of a new salt bridge between Glu317 and Arg171 from helix 1. However, R171A mutation had little effect on the structure of either the unliganded or the peptide-bound state from which we infer that the Glu317-Arg171 salt bridge is not crucial for the binding of collagen or α1I activation. In contrast, mutation R287A or E317A produced significant conformational change in the α1I domain. Analysis of the NMR data provided strong evidence that helix 7 was displaced in both of these mutants, suggesting that the salt bridge observed in the unliganded structure is important in maintaining the low affinity conformation. However, the two mutants showed different patterns of chemical shift perturbations in NMR spectra for residues at the collagen-binding site, suggesting that they have differential effects on the structure of binding site. These differences may explain the different effects of these mutations on the collagen binding affinity observed. During the course of the investigation, we also identified and characterized a dimeric complex of α1I-GLOGEN that supports the binding pose determined for the monomer. These data together provided a more complete structural profile of the collagen-bound α1I domain that advances our understanding of signaling and specificity in collagen-binding integrins.
Structural Analysis of Integrin αII Domain-Collagen Complex

EXPERIMENTAL PROCEDURES

Production of αII Domain

Uniformly $^{15}$N-labeled, $^{15}$N,$^{13}$C-labeled, and $^2$H,$^{15}$N,$^{13}$C-labeled human αII was prepared as described previously (31). For expression of unlabeled αII, the same protocol was followed except Luria broth (Sigma-Aldrich) was used instead of isotopically labeled minimal medium, and expression was initiated by induction with isopropyl 1-thio-β-d-galactopyranoside (Astral Scientific, Australia; 1 mM) at an $A_{600}$ of 0.6.

Synthesis of Collagen Peptide

A collagen peptide containing the sequence Ac-GPOGPO-GLOGENGPOGPOGPO-NH$_2$ was synthesized and purified as described previously (32) except that Fmoc (N-(9-fluorenyl)methoxy-carbonyl) Rink-amide resin (Chem-Impex International) was used. The peptide is referred to hereafter as GLOGEN.

Backbone Resonance Assignment

NMR experiments were conducted at 298 K on a Bruker Avance 800 spectrometer equipped with a cryoprobe. The sample contained $^{15}$N,$^{13}$C-labeled αII (1.2 mM) and unlabeled GLOGEN peptide (2.4 mM) in the NMR buffer (50 mM HEPES, pH 7.4, 50 mM NaCl, 5 mM MgCl$_2$, and 10% $^2$H$_2$O). The chemical shifts were referenced according to the method described by Wishart et al. (33) in which the $^1$H chemical shifts were referenced to the water peak, whereas the $^{15}$N and $^{13}$C chemical shifts were referenced by the $^{15}$N/$^1$H and $^{13}$C/$^1$H gyromagnetic ratios. For the backbone resonance assignments, two-dimensional $^1$H,$^{15}$N HSQC, three-dimensional HNCA, three-dimensional HNCACB, and three-dimensional HNCO spectra were acquired. Except where noted otherwise, spectra were processed using Topspin (version 3.0, Bruker-BioSpin™), and the analysis was carried out using SPARKY (34).

Titration of Wild-type (WT) αII with GLOGEN Using Two-dimensional NMR

To observe GLOGEN binding to αII, a series of two-dimensional $^1$H,$^{15}$N SOFAST HMQC (35) spectra on uniformly $^{15}$N-labeled αII was acquired with stepwise addition of peptide. The samples contained αII (100 μM) in NMR buffer. GLOGEN was added from a stock solution (10 mM in water) to give final peptide concentrations of 22.5, 45, 90, 180, and 360 μM. Spectra were recorded at 293 K on a Bruker Avance 600-MHz NMR spectrometer equipped with a cryoprobe. Chemical shift perturbations (CSPs) caused by peptide binding were measured from two-dimensional $^1$H,$^{15}$N TROSY spectra of uniformly $^2$H,$^{13}$C,$^{15}$N-labeled αII (400 μM) in NMR buffer in the presence and absence of GLOGEN (800 μM). Spectra were recorded at 293 K on a Varian INOVA 600-MHz NMR spectrometer fitted with a cryogenically cooled probe. CSPs were derived for each assigned residue using the following equation.

\[
CSP = \sqrt{0.154 \times \Delta N^2 + \Delta H^2} \quad \text{(Eq. 1)}
\]

Analysis of αII Mutants Using Two-dimensional NMR

$^1$H,$^{15}$N TROSY spectra on uniformly $^{15}$N-labeled R171A, R287A, and E317A αII mutants were acquired. The samples contained 275 μM R171A, 200 μM R287A, or 100 μM E317A in NMR buffer. To observe the binding of GLOGEN, the peptide was added to a final concentration of 560 μM in the case of R171A, 200 μM in the case of E317A, or 560 μM in the case of R171A. Spectra were recorded at 293 K on a Bruker Avance 600-MHz NMR spectrometer equipped with a cryoprobe.

T1 and T2 Relaxation

T1 and T2 relaxation data were acquired for WT αII as described by Farrow et al. (36) at 293 K on a Varian INOVA 600-MHz NMR spectrometer equipped with a cryogenically cooled probe. The unliganded sample contained $^{15}$N-labeled αII (400 μM). The peptide-bound sample contained $^{15}$N,$^{13}$C-labeled αII (1.2 mM) and GLOGEN (2.4 mM). Samples were prepared in the NMR buffer. The relaxation delay was sampled at 0.05, 0.1, 0.2, 0.4, 0.8, 1.2, 1.6, and 2.0 s for the longitudinal relaxation measurement and 0.01, 0.03, 0.07, 0.09, 0.11, 0.13, and 0.15 s for the transverse relaxation measurements. A two-dimensional heteronuclear $^1$H,$^{15}$N NOE spectrum was acquired with 3 s of weak irradiation at either the center of the amide proton frequency range to generate the heteronuclear NOE or 10,000-Hz off-resonance for the no-NOE control. Spectra were processed using NMRPipe (37), and the signal decay was analyzed and plotted using SPARKY (34). Theoretical T1/T2 ratios for different structural models were generated using the program HYDRONMR (38).

Side-chain Assignments and NMR Structure Calculation

For the side-chain assignments and generation of structural restraints, three-dimensional HBHA(CBCA(CO)NH) was recorded on a Bruker Avance 500-MHz spectrometer fitted with a cryoprobe, three-dimensional $^1$H,$^{15}$N NOESY,$^{15}$N HSQC and three-dimensional $^1$H,$^{13}$C HSQC (aromatic) experiments were recorded on a Bruker Avance 800-MHz spectrometer fitted with a cryoprobe, and three-dimensional $^1$H,$^{13}$C HSQC (aliphatic) was recorded on a Bruker Avance 900-MHz spectrometer fitted with a cryoprobe. All spectra were recorded at 298 K. For the NOESY spectra, the mixing time was 60 ms, and the $^{13}$C-carrier frequency was 28 and 123.5 ppm for the aliphatic and aromatic $^{13}$C-edited NOESY spectra, respectively. CARA (Computer-aided Resonance Assignment) was used for spectral analysis. Automated structure calculation was performed using the software package UNIO-ATNOS/CANDID (39, 40) in combination with the torsion angle dynamics program CYANA 3.0 (41) following the protocol described by Serrano et al. (42). Structures were calculated initially in the absence of a Mg$^{2+}$ ion. The 20 conformers with the lowest target function values were analyzed to determine the site of Mg$^{2+}$ ion binding. Based on the CSP pattern of αII on Mg$^{2+}$ ion binding and the crystal structure of α2I-collagen peptide complex, the Mg$^{2+}$ was assigned to coordinate with residues Ser$^{152}$, Ser$^{154}$, and Thr$^{220}$ in our αII structures. Subsequently, a Mg$^{2+}$ ion was introduced to the structures using a “pseudo-link” consisting of 30 “pseudo-residues” extending from the C terminus. Structures were recalculated in the presence of the Mg$^{2+}$ ion using UNIO-ATNOS/CANDID as described above with the addition of backbone torsion angle restraints generated by the program TALOS+ (43). In addition,
hydrogen bond restraints (two per bond) were added in regions of canonical secondary structure where a unique hydrogen bond donor-acceptor pair was evident from structural convergence. The 20 structures having the lowest target function with Mg$^{2+}$ bound were then refined using Cartesian dynamics in CNS (44) after removal of the pseudo-link to produce the final structures. The 20 lowest energy conformers with no NOE violations >0.25 Å, no bond violations >0.05 Å, and no improper or dihedral angle violations >5° were chosen to represent the solution structure of liganded αI1.

**HADDOCK Docking**

The docking of collagen peptide to αI1 was performed using the data-driven docking program HADDOCK (45, 46).

**αI1 Template**—The 20 lowest energy structures of αI1 in its liganded state were used as the protein input templates. Restraints that were used in the original structure calculation, i.e. NOEs, H-bonds, and dihedral angle restraints, were included in the docking process to constrain the protein in its liganded conformation. Regions of the protein with low T1/T2 ratios, low values of the heteronuclear NOE, and low angular order ($S^2$) values predicted from chemical shifts were allowed to be fully flexible during docking. HADDOCK requires a set of ambiguous interaction restraints (AIRs) at the binding interface that are divided into “active” and “passive” categories where active residues are those directly implicated in binding from experimental data and passive residues are their near neighbors. Residues on αI1 for which active AIRs were generated were selected based on their having NMR CSPs >0.5 ppm and a loss of NMR signal intensity >70% upon titration of αI1 with GLOGEN. Exclusion criteria were residues in regions of the protein that showed fast time scale dynamics in the heteronuclear NOE experiment and residues that were not solvent-accessible.

**GLOGEN Peptide Template**—For the peptide input template, the initial model was obtained by computationally modifying an available crystal structure of a triple-helical collagen peptide (Protein Data Bank code 1Q7D) (47). The peptide in the crystal structure contains a GFOGER recognition motif, which was converted to GLOGEN using the mutagenesis function in PyMOL (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC). Although the peptide is a homotrimer, each of the strands is in a unique chemical environment, referred to as the “leading,” “middle,” and “trailing” strands, as viewed from their N termini. To make sure that the triple-helical structure was maintained during the docking, hydrogen bonds present in the crystal structure were included as input restraints. AIRs were also generated for the peptide based on previous studies that showed the essentiality of the glutamate residue in the GLOGEN motif. It was known that one of the three glutamate residues is responsible for coordinating with the metal ion at the αI1, but it was unclear whether any one of the glutamate residues in the trimeric peptide model is preferred over the other two. To predict whether αI1 has a preference for coordinating to any one of the three strands, a preliminary docking was conducted with an ambiguous distance restraint set such that any of the three glutamate residues of the triple-helical peptide may coordinate with the magnesium ion. Residues within 8 Å of a glutamate were defined as active, and residues between 8 and 12 Å from the glutamate were defined as passive. This length matches the diameter of the binding interface of αI1 and is likely to cover all potential residues on the peptide that could make direct interactions with αI1. Based on the lowest energy cluster arising from the preliminary dock, a unique peptide interface was defined, and seven active and eight passive residues were selected as the “optimized” peptide AIRs. The active residues comprised Hyp$^1$, Gly$^3$, and Asn$^{12}$ from the leading strand and Hyp$^{110}$, Gly$^{111}$, Asn$^{113}$, and Gly$^{114}$ from the middle strand. The passive residues comprised Gly$^8$, Leu$^9$ from the leading strand; Gly$^{108}$, Leu$^{109}$, Pro$^{115}$, and Hyp$^{116}$ from the middle strand; and Pro$^{215}$ and Pro$^{216}$ from the trailing strand. The leading strand glutamate-magnesium ion interaction was assigned as an unambiguous distance restraint as previous studies have confirmed its essentiality for binding (10). The docking process included a rigid body energy minimization step, which produced 1000 structures. The best 200 structures were subjected to a semiflexible simulated annealing step and then a final low temperature flexible refinement in explicit waters.

**Mutagenesis**

Mutagenesis was performed using the QuikChange II mutagenesis kit (Agilent Technologies) according to the manufacturer’s instructions. The pET28a plasmid containing the WT αI1 insert was used as the template with the following primers used to generate the R171A and E317A mutations: R171A: forward, 5'-AGCTTTTTTAAATGACCTTCTTGAAGCAATGGATAT-TGGTCCTAAACAGACA-3'; reverse, 5'-TGCTCTTTAGGACCAATATCCATTGCTTCAAGAAGGTCATTTAAAAA-GCT-3'; E317A: forward, 5'-AAGCATTTCTCAATGGTCTGTTAGGACCAATATCCATTGCTTCAAGAAGGTCATTTAAAAA-GCT-3'; reverse, 5'-TTTAAATGTTGACTAGACCGAAAGGATCGACAAATTGAAGAAAATGCTT-3'. The presence of mutations was verified by DNA sequencing. For the R287A mutant, a synthetic gene encoding αI1 with the mutation in the pET28a plasmid was ordered from DNA2.0. The mutants were expressed in 15N-labeled minimal medium following the same protocol used for the WT αI1 integrin.

**Analytical Size Exclusion Chromatography (SEC)**

Analytical SEC was carried out using a Superdex 75 HR 10/30 column (GE Healthcare) with a bed volume of 23.6 ml on an AKTA$^\text{TM}$ purifier protein chromatography system. Samples (100 μl) containing αI1 (10 μM) with or without GLOGEN (20 μM) were applied to the column equilibrated with 50 mM HEPES buffer, pH 7.4, 50 mM NaCl, and 5 mM MgCl$_2$. A flow rate of 0.7 ml/min was maintained, and the elution was monitored by continuous measurement of UV absorbance at 280 nm ($A_{280}$). The runs were conducted at room temperature.

**SEC-SAXS**

All SAXS data were acquired at the Australian Synchrotron SAXS/WAXS beamline and the data collection and scattering-derivative parameters are described in Table 2 according to the recommendations of the International Union of Crystallography Commission on Small-Angle Scattering (48). The SAXS experiments were set up in line with gel filtration chromatography as described by Gunn et al. (49). Samples (50 μl) contain-
Structural Analysis of Integrin α1I Domain-Collagen Complex

RESULTS

GLOGEN Binding to α1I Observed by NMR—The sequence of the peptide used in the study contained the integrin recognition sequence GLOGEN flanked by five GPO repeats. Its ability to self-assemble into the collagen-like triple-helical conformation was confirmed using circular dichroism (data not shown). The GLOGEN motif is the most potent ligand for α1β1 among the fibrillar collagen sequences, and it inhibits the binding of α1β1 to collagen type IV with an IC50 of ~3 µM (21). Comparison of two-dimensional 1H,15N TROSY spectra of the α1I domain in the presence and absence of the GLOGEN peptide revealed extensive CSPs upon addition of excess peptide (Fig. 1). Gradual titration of α1I with the peptide revealed that at a low GLOGEN:α1I ratio (<1:1) the two-dimensional 1H,15N SOFAST HMQC spectra of α1I exhibited severe line broadening for most peaks in the spectrum (Fig. 1C). The intensity of the peaks increased as the peptide concentration exceeded that of α1I, and at a GLOGEN:α1I ratio of ~3:1, most of the signals were recovered (Fig. 1D). Similar broadening has previously been reported upon titration of the α2I domain with a peptide containing the GFOGER recognition motif (54). Such broadening

FIGURE 1. HSQC spectra of the titration of α1I with GLOGEN peptide. A, overlay of the 1H,15N TROSY of unliganded α1I (red) and a 1:2 complex of α1I bound to GLOGEN (black) showing significant chemical shift perturbations. The samples contained 2H,15N,13C-labeled α1I (400 µM) in the presence and absence of GLOGEN (800 µM). The spectra were acquired at 600 MHz and 293 K. B–D, 1H,15N SOFAST HMQC spectra of 1H,15N-labeled α1I (100 µM) with no GLOGEN (B) and GLOGEN:α1I at a ratio of 1:1 (C) and 3:1 (D). NMR spectra were acquired at 800 MHz and 293 K. The spectral quality declined during the early stage of GLOGEN titration but was gradually recovered as excess peptide was titrated into the system.
of NMR resonances is often the result of intermediate time scale chemical exchange. However, as the homotrimeric collagen-mimetic peptides potentially have three αI domain binding sites, the formation of a larger oligomeric complex at low peptide:αI ratios is a possible explanation for the loss of signal intensity.

**T1/T2 Analysis of the Stoichiometry of the Complex**—The oligomeric state of the αI-GLOGEN complex was assessed using NMR by measuring the heteronuclear 1H,15N T1/T2 relaxation ratios for αII domain in the absence and presence of a 2-fold excess of GLOGEN. The experimental data were compared with theoretical data generated using HYDRONMR (38) for the isolated αI domain (Protein Data Bank code 1QCY (55)), isolated αII domain (Protein Data Bank code 1AOX (56)), the monomeric GFOGER-αII complex (Protein Data Bank code 1DZI (22)), and a model that consisted of two αII domains bound to one triple-helical peptide (Fig. 2). The model of the 2:1 complex was generated manually by docking a second αII molecule onto the structure of the GFOGER-αII complex. The theoretical T1/T2 ratios for the 2:1 complex were significantly higher than the rest of the models due to its size and shape. The fluctuation in the T1/T2 values is related to the orientation of the amide N–H bond vectors with respect to the principle axis frame of the rotational diffusion tensor in the calculation. This kind of variation is typical for elongated ellipsoidal proteins such as that of the 2:1 complex. The global T1/T2 ratio measured for the αII-GLOGEN complex is in good agreement with the theoretical data derived for the monomeric complex and is consistent with the formation of a 1:1 complex between αII domain and the triple-helical GLOGEN at high GLOGEN:αII domain ratios.

**Structure of the αII Domain Bound to GLOGEN**—The structure of αII domain bound to GLOGEN was determined using a sample containing 15N,13C-labeled αII (1.2 mM) in the presence of Mg2+ (5 mM) and unlabeled GLOGEN (2.4 mM). The pattern of CSPs observed upon addition of Mg2+ to αII was consistent with the position of the metal ion in the crystal structure of GFOGER-αII complex (31). The coordination of Mg2+ in our structure was therefore assumed to be similar to the αII complex structure. An ensemble of 20 conformers representing the structure of αII bound to GLOGEN (from a total of 40 structures calculated) is shown in Fig. 3. Structural statistics for the ensemble are shown in Table 1. The structure assumes a typical Rossmann fold consisting of five parallel β-strands and one antiparallel β-strand surrounded by six α-helices. Comparison of the lowest energy model with the unliganded structure of αII (Fig. 4) shows that the central β-sheet core remains sim-
Structural Analysis of Integrin α11 Domain-Collagen Complex

FIGURE 4. Effects of GLOGEN binding on α11. A, TALOS+–predicted $S^2$ value for the unliganded (red) and GLOGEN-bound (black) α11 based on the NMR chemical shifts. $S^2$, the angular order parameter, describes the rigidity of a residue as a measure of its backbone conformational entropy. It ranges between 0 for a completely disordered residue and 1 for a rigid residue. Significant drops at the βE-α6 loop (residues 281–292). Very few NOEs could be assigned in the βE-α6 loop in NOESY spectra of the complex. The TALOS+–predicted angular order ($S^2$) values (43) showed a clear decrease in the predicted order in the βE-α6 loop of α11 upon binding to GLOGEN (Fig. 4A). This is supported by comparison of the $^1$H,${^15}$N heteronuclear NOE for α11 unliganded and in complex with GLOGEN that revealed a decrease upon peptide binding, consistent with increased dynamics in this region in the bound state (Fig. 4B). These findings are consistent with the lower precision observed in this region of the structure of α11 (Fig. 3A).

The second region undergoing major conformational change is the C-terminal helix (helix 7), which is displaced downward by 12 Å upon GLOGEN binding relative to its position in the unliganded structure. In the unliganded state, helix 7 is linked to the C helix by a salt bridge between Arg287 and Arg317. This salt bridge is broken in the complex with GLOGEN. In its new position in the complex, Glu317 is in a position from where it may form a salt bridge with Arg177 from helix 1 (Fig. 4F). To investigate the importance of the salt bridges in regulating the conformational changes, single point mutants with R171A, R287A, and E317A substitutions were studied (Fig. 5).

The $^1$H,${^15}$N TROSY spectrum of the R171A mutant (275 μM) in the absence and presence of GLOGEN (500 μM) showed a high degree of similarity to the spectrum of the unliganded and peptide-bound states of WT α11, respectively. Chemical shift changes were only observed for residues located in close proximity to the mutation site in each case (Fig. 5, A and B), suggesting that the mutation had little impact on the structure of either the unliganded or peptide-bound state of α11.

In contrast, the $^1$H,${^15}$N TROSY spectrum of R287A (200 μM) in the absence of GLOGEN showed substantial differences from the spectrum of the unliganded WT α11. Comparison of the spectrum with the GLOGEN-bound WT α11, however, showed a much higher degree of similarity (Fig. 5, C and D). In particular, a strong peak from the spectrum of the unliganded R287A coincides with the peak corresponding to the C-terminal residue, Ile331, in the GLOGEN-bound WT α11 (Fig. 5G). The chemical shift and intensity of this peak suggest that this similar with an root mean square deviation on the Ca of 1.1 Å. Major conformational differences were identified in two regions. The first region is the C helix of the unliganded state that becomes completely unfolded upon peptide binding. Unwinding of the short C helix results in extending the βE-α6 loop (residues 281–292). Very few NOEs could be assigned in the βE-α6 loop in NOESY spectra of the complex. The TALOS+–predicted angular order ($S^2$) values (43) showed a clear decrease in the predicted order in the βE-α6 loop of α11 upon binding to GLOGEN (Fig. 4A). This is supported by comparison of the $^1$H,${^15}$N heteronuclear NOE for α11 unliganded and in complex with GLOGEN that revealed a decrease upon peptide binding, consistent with increased dynamics in this region in the bound state (Fig. 4B). These findings are consistent with the lower precision observed in this region of the structure of α11 (Fig. 3A).

The second region undergoing major conformational change is the C-terminal helix (helix 7), which is displaced downward by 12 Å upon GLOGEN binding relative to its position in the unliganded structure. In the unliganded state, helix 7 is linked to the C helix by a salt bridge between Arg287 and Arg317. This salt bridge is broken in the complex with GLOGEN. In its new position in the complex, Glu317 is in a position from where it may form a salt bridge with Arg177 from helix 1 (Fig. 4F). To investigate the importance of the salt bridges in regulating the conformational changes, single point mutants with R171A, R287A, and E317A substitutions were studied (Fig. 5).

The $^1$H,${^15}$N TROSY spectrum of the R171A mutant (275 μM) in the absence and presence of GLOGEN (500 μM) showed a high degree of similarity to the spectrum of the unliganded and peptide-bound states of WT α11, respectively. Chemical shift changes were only observed for residues located in close proximity to the mutation site in each case (Fig. 5, A and B), suggesting that the mutation had little impact on the structure of either the unliganded or peptide-bound state of α11.

In contrast, the $^1$H,${^15}$N TROSY spectrum of R287A (200 μM) in the absence of GLOGEN showed substantial differences from the spectrum of the unliganded WT α11. Comparison of the spectrum with the GLOGEN-bound WT α11, however, showed a much higher degree of similarity (Fig. 5, C and D). In particular, a strong peak from the spectrum of the unliganded R287A coincides with the peak corresponding to the C-terminal residue, Ile331, in the GLOGEN-bound WT α11 (Fig. 5G). The chemical shift and intensity of this peak suggest that this GLOGEN-bound α11 (black). The low values for residues in the βE-α6 loop and at the C terminus are consistent with increased flexibility in the presence of GLOGEN. C, percentage of signal intensity loss in the $^1$H,${^15}$N TROSY spectrum as a consequence of peptide binding. Residues with cross-peaks that lost more than 70% intensity are highlighted in red. D, weighted chemical shift perturbations measured in the TROSY spectra highlighted regions (highlighted with gray bars across the other panels) that were most significantly influenced by peptide binding. Regions with the largest CSPs include the three MIDAS loops, βE-α6 loop, and helix 7, which is distant from the MIDAS but for which a 12 Å downward displacement was observed in the structure. E and F, aligned structures with these regions highlighted for the unliganded (red; Protein Data Bank code 1QCY) and GLOGEN-bound (black) α11 domain. E shows a zoomed image of the C helix, which is uncoiled in the liganded structure. The spheres represent the Mg$^{2+}$ ions. In F, the salt bridge formed between residue Glu317 and Arg287 in the unliganded (red) or Arg177 in the liganded (black) states is indicated as a black dotted line.
FIGURE 5. NMR analysis of α11 mutants. Overlaid spectra of 1H,15N-labeled WT (red) and R171A (blue) α11 (A) and WT α11 (black) and R171A (magenta) α11 (B) both in the presence of GLOGEN are shown. The similarity of the WT and R171A spectra in both their unliganded and GLOGEN-bound states suggests that the R171A mutation had little effect on peptide binding. C, unliganded WT (red) and R287A (blue) α11. D, WT α11 with GLOGEN (black) and unliganded R287A α11 (blue). The spectrum of the unliganded R287A mutant showed a greater degree of similarity to the peptide-bound state of WT α11, suggesting that the unbound state of the R287A mutant adopts a conformation that is more similar to the activated state of the WT α11. E, unliganded WT (red) and E317A (blue) α11. F, WT α11 with GLOGEN (black) and unliganded E317A α11 (blue). The spectrum of the unliganded E317A was also more similar to the spectrum of the peptide-bound state of WT α11, suggesting that the E317A also adopts a similar conformation. G and H, zoomed regions of the TROSY spectra highlighting residue Ile331 in unliganded WT α11 (red), GLOGEN-bound WT α11 (black), and unliganded R287A (blue) (G) or E317A (blue) (H). NMR spectra were acquired at 800 MHz and 293 K.
cross-peak in the spectrum of the R287A mutant also represents the C-terminal residue, implying that helix 7 is displaced in the R287A mutant in the absence of peptide. Addition of GLOGEN (400 μM) to R287A induced relatively minor chemical shift changes, and the spectrum also resembles the spectrum of the GLOGEN-bound WT α11. With the aid of the 15N,1H assignments of the GLOGEN-bound WT α11, residues that were more severely perturbed by the R287A mutation in the bound state were identified. The majority of perturbed residues are located in close proximity either to the site of mutation or to residues within or adjacent to the MIDAS (Fig. 6).

As was the case with R287A, the 1H,15N TROSY spectrum of E317A in the absence of GLOGEN displayed significant differences from the spectrum of unliganded WT α11. The spectrum appeared to be more similar to the GLOGEN-bound state of WT α11 (Fig. 5, E and F). The spectrum of E317A also contains a strong peak at a chemical shift that is coincident with the C-terminal residue (Ile331) in the GLOGEN-bound spectrum of WT α11, suggesting that helix 7 in E317A is displaced (Fig. 5H). Addition of GLOGEN (400 μM) to E317A resulted in relatively minor changes to the spectrum. However, the pattern of perturbations relative to WT α11 was different from that observed with R287A. Fewer perturbations were observed for residues adjacent to the MIDAS in the spectrum of E317A. In contrast, the majority of the perturbed residues were located in helix 1, βE-α6 (the uncoiled C helix), and βF sheet (Fig. 6).

Mapping the Binding Interface—The CSPs and chemical exchange broadening that were observed in the two-dimensional 1H,15N TROSY spectrum of WT α11 upon GLOGEN binding were assessed to map the binding interface of the complex. CSPs and chemical exchange broadening are observed for residues that undergo a change in their chemical environment upon binding. This can be caused both by direct interaction with the ligand and indirect effects such as those due to conformational change. The analysis revealed that 22 residues lost more than 70% of their signal intensity on peptide binding. These residues were mostly located in the three MIDAS loops, loop βE-α6, and helix 6. CSPs on the other hand mapped both to the same regions and to residues in helix 7 (Fig. 4, C and D).

Structure of the α11-GLOGEN Complex—A structure of the α11-GLOGEN complex was generated using the data-driven docking program HADDOCK (45, 46) (Protein Data Bank code 2M32). Based on the selection criteria described above, residues in helix 7 were excluded from the definition of the binding interface as they showed no broadening, and it was inferred that the CSPs were due to the observed conformational change in this region rather than a direct interaction with GLOGEN. In addition, residues 281–292 from the βE-α6 loop were excluded from the binding interface as they showed fast time scale dynamics in the heteronuclear NOE experiment and were therefore deemed to be disordered in the bound state. The residues of α11 defined as active for the HADDOCK calculation comprised the surface-exposed residues with CSP >0.5 ppm as well as residues with a loss of signal intensity >70%, namely Ser154, Ile155, Tyr156, Arg218, Gln219, Glu255, and His257. Only Asn153 satisfied the criteria for being a passive residue. These AIRs were used as input for the HADDOCK calculation. No intermolecular NOEs were included in the HADDOCK calculations. Analysis of the 200 final structures using their HADDOCK score (45, 46) produced a cluster of 10 structures, all of which fell within the 20 best scoring structures. This cluster was selected to represent the structure of triple-helical GLOGEN peptide bound to α11 (Fig. 7). The lowest energy structure of the 10 was selected to best represent the model of the complex. It shows the peptide binding along a “trench” above the MIDAS (Fig. 7C). Consistent with the α2I-GFOGER complex crystal structure (22), the α11 molecule binds exclusively to the leading (cyan) and middle (orange) strands as shown in Fig. 7F. The trailing strand (green) does not make any interaction with α11. The binding orientation of the peptide, however, is different from the α2I crystal structure in which GFGER sits along the edge of the trench (Fig. 7, compare C and D) (22). A more detailed discussion and comparison of the two structures is presented below.

Analysis of the Complex at Low GLOGEN:α11 Ratios—Potential causes of the broadening observed in the 1H,15N TROSY spectrum of α11 at lower peptide concentrations were investigated using analytical SEC and SAXS. Analysis of the SEC elution profile of a sample containing α11 and GLOGEN peptide revealed the presence of two distinct peaks. The first of these eluted at a volume similar to unliganded α11 (11.62 ml), whereas the second, which was inferred to correspond to the α11-peptide complex, eluted considerably earlier (10.02 ml) (Fig. 8). The >1.5-ml shorter elution volume suggested that the molecular weight of the complex species was much higher than that of the unliganded state. Comparison of the elution volume with other proteins with different molecular weight on the same column3 suggested that this peak corresponds to a species that is much bigger than the expected 1:1 GLOGEN-bound α11 species (molecular mass, 27.4 kDa).

SAXS experiments were set up in line with SEC. The sample containing the mixture of α11 and GLOGEN eluted earlier than the unliganded α11 sample, indicating a larger species in the mix-

3 The typical elution time of albumin and ovalbumin from a gel filtration column (Superdex 75 HR 10/30 column, GE Healthcare) was published by European Molecular Biology Laboratory (EMBL) on the website www.embl.de/pepcore/papcore_services/protein_purification/chromatography/superdex75HR10-30.
Porod analysis of the scattering data indicated that the estimated molecular mass of α1I-GLOGEN complex was 40.7 ± 0.5 kDa, whereas that of the unliganded α1I was 19.2 ± 0.2 kDa. Although the SAXS-derived parameters such as $R_g$ and volume were stable across much of the peak (varying by less than 2%) and consistent with a complex containing two α1I molecules, the resolution of the small bed volume column used was insufficient to cleanly separate the larger species from later eluting species such as 1:1 α1I-GLOGEN complex or free α1I. Indeed, single value decomposition analysis using SvdPlot in PRIMUS (53) of the entire elution consisting of 11 SAXS data sets from across the peak suggested a mixture of at least three species.

The $R_g$ values, distance distribution functions $P(r)$, and calculated maximum dimension ($D_{max}$) (Fig. 9) also suggested the presence of a GLOGEN-bound α1I species that was considerably bigger and more elongated than the unliganded α1I or a putative 1:1 α1I-GLOGEN complex. Therefore, we postulated that a complex composed of two α1I molecules bound to a single homotrimeric peptide may have formed. The α1I domain can in theory bind to any two of the three strands of the peptide to form three different combinations of two α1I domains bound to one GLOGEN triple helix (2:1 complex). These three models were built using the lowest energy HADDOCK model as a template. Two of the three combinations showed significant steric clashes and were excluded from further investigation (data not shown). Although the SAXS data were likely to be from a mixture, the Porod mass estimate suggested that the volume fraction of any smaller species was reasonably small (≈10%). Therefore, ab initio reconstruction from the SAXS was performed, and an average-filtered shape envelope was generated. Although the normalized special discrepancy of the shape envelopes was low (0.711 ± 0.067), a relatively high reduced χ² statistic (χ²) of 2.13 ± 0.02 was noted from the DAMMIF reconstructions. This appears to be largely due to the detector count statistics of the synchrotron SAXS data underestimating the true uncertainty but may also reflect the inability of a single shape to adequately describe the scattering from the mixture.
The HADDOCK-derived 2:1 model shows good shape correspondence with the SAXS envelope (Fig. 9). Despite the overall shape similarity between the HADDOCK-derived model and average-filtered shape (normalized special discrepancy of 0.94), the fit of the theoretical scattering profile from this model to the SAXS data generated using CRYSOl (51) was relatively poor with a $\chi^2$ of 5.3. There are several plausible reasons why this is the case. Again, the issues of detector count statistics and the presence of other species (e.g. free peptide, unbound $\alpha$II, and monomeric complex) would have contributed to the high value. In addition, the GPO repeats at the tails of the GLOGEN peptides from the HADDOCK model fit particularly poorly in the envelope, suggesting that they may be flexible. To address the issue of mixed species, OLIGOMER (53) was used to estimate the scattering contribution of the different species present. The best fit obtained was for 88% 2:1 complex, 7% 1:1 complex, and 5% free peptide. The $\chi^2$ for this volume fraction composition was 3.1, which represents a statistically significant improvement in the goodness of fit ($F = 1.71; P_F < 3 \times 10^{-7}$) over the 2:1 complex alone or any other combination of species tested. It seems likely that further improvements in goodness of fit could be obtained by considering flexibility either in the GLOGEN peptide tails or in the $\beta\epsilon-\alpha$6 loop and helix 7 and/or some variability in the relative $\alpha$II domain orientation in the 2:1 complex. Nonetheless, in relative terms, the HADDOCK-derived 2:1 model provides a statistically significant improvement in fit compared with any other 2:1 model tested including the corresponding 2:1 complex based on the GFOGER-$\alpha$2I crystal structure, suggesting that the difference in peptide binding orientation observed between the $\alpha$II and $\alpha$2I complex is genuine.

**DISCUSSION**

We have determined the structure of the $\alpha$II domain bound to a triple-helical peptide that is a mimic of its natural ligand.
collagen. The structure reveals that binding is accompanied by movement of the MIDAS loops as well as major conformational changes in the C helix (residues Leu282–Gly288), which became uncoiled in the bound state, and a displacement of the C-terminal helix 7 (helix 7), which moved by ∼12 Å relative to its position in the unliganded state. These structural rearrangements are consistent with those observed in the structure of the bound state using hydrogen-deuterium exchange mass spectroscopy (57). The study provided evidence that helix 7 is displaced in both cases. However, it is also apparent from analysis of the spectra that the two mutations have different effects on residues at the collagen-binding site of the unliganded and liganded states using $^{1}H,^{15}N$ TROSY NMR.

Our results do not support this interpretation. We have acquired $^{1}H,^{15}N$ TROSY for both mutants, and with the benefit of full assignments for WT α1I, the spectra provide strong evidence that helix 7 is displaced in both cases. However, it is also apparent from analysis of the spectra that the two mutations have different effects on residues at the collagen-binding site of α1I (Fig. 6). Thus, we believe that the differences observed in the previous study may have resulted from the mutations altering the collagen binding affinity of the two proteins.

The HADDOCK model of α1I in complex with GLOGEN shows that the peptide binds along a surface trench above the MIDAS on α1I (Fig. 7C). In contrast, in the structure of α2I in complex with GLOGER, the peptide binds at the edge of a similar trench (Fig. 7D) (22). Our SAXS data support the HADDOCK model of α1I-GLOGEN, whereas a model based on the α2I-GLOGER crystal structure showed a very poor fit to the SAXS envelope. Analysis of the binding interfaces provides some indications as to the structural basis of these differences. First, the asparagine residues of the leading strand GLOGEN motif sits in an acidic pocket formed by Asp253 and Gly254 of α1I (Fig. 7E). Replacing the asparagine with a bulkier arginine could potentially create a steric clash in that pocket, which may...

**Table 2**

| Data geometry parameters | Instrument | Australian Synchrotron SAXS/WAXS beamline with inline size exclusion chromatography |
|--------------------------|------------|-------------------------------------------------------------------------------------|
| Beam geometry            | Point collimation (250 μm horizontal × 150 μm vertical) |
| Wavelength (Å)           | 1.0322     |
| Samples                  | apo α1I WT, α1I WT + GLOGEN |
| g range (Å⁻¹)            | 0.030–0.444 |
| Exposure time (s)        | 45 × 2     |
| Concentration range (mg mL⁻¹) | 0.4–1.2 |
| Temperature (K)          | 298        |

Structural parameters:
- $R_{g}$ (Å) from Porod: 15.5 ± 0.02
- $R_{g}$ (Å) from Guinier: 18.0 ± 0.1
- Porod volume estimate (Å³): 370 ± 300
- Dry volume calculated from sequence (Å³): 26000

Molecular-mass determination:
- Molecular mass Mr (kDa): 19.2 ± 0.2
- Porod volume (kDa): 21.5
- Calculated monomeric Mr from sequence (kDa): 27.4

Software employed:
- Primary data reduction: SASS15ID
- Data processing: AUTOPODOR
- Ab initio analysis: DAMMIF
- Validation and averaging: DAMAVER
- Rigidity-body modeling: CRYOSOL
- Volume fraction analysis: OLGOMER
- Three-dimensional graphics representations: PyMOL

+ Reported for the averaged SAXS data where Porod volume estimates from successive blocks of 5 × 2-s exposures did not vary by more than 1.5% from the mean value.

+ Calculated for two α1 monomers bound to a GLOGEN triple helix.

N/A, not applicable.
account for the observation that GLOGEN binds more potently to the α1I domain than GFOGER. Second, comparison of the sequences indicates a change where residue Ala<sup>218</sup> in αII corresponds to Asp<sup>219</sup> in α2I. In the structure of the α2I-GFOGER complex, Asp<sup>219</sup> is located at the edge of the trench and forms a salt bridge with one of the arginine residues in the GFOGER motif. The change to Ala<sup>218</sup> in αII precludes the formation of a similar interaction in the αII complex, which may contribute to both the peptide binding preferences observed for the two αI domains (15, 21) and the subtle change in binding orientation of the peptides between the two complexes.

Homotrimeric helical peptides have been used widely as collagen mimetics and represent excellent tools to study the binding of collagen on collagen receptors. They have been especially helpful in the process of identifying specific recognition motifs for collagen-binding integrins (17). For such homotrimeric peptides, there are potentially three equivalent binding sites that may support binding to integrin αI domains. Therefore, in a case where the individual strands in the peptide are in the appropriate orientation for αI binding and no steric hindrance is present, more than one strand in the same peptide should in theory be able to bind to an individual αI molecule simultaneously. Herein, our size exclusion and SAXS results provide evidence for the existence of a dimeric complex for αI1 bound to GLOGEN. To our knowledge, this is the first report describing a dimeric complex between an αI domain and a triple-helical collagen peptide, but we believe that the SEC result presented by Lambert et al. (54) may also demonstrate the same phenomenon in α2I upon peptide binding, although it was explained as an artifact resulting from the elution of the rigid, elongated helical peptide. In the case of αII1 bound to GLOGEN, T1/T2 NMR relaxation data confirmed the formation of a 1:1 monomeric complex at higher peptide concentrations. Based on these observations, we propose that in the presence of GLOGEN αII1 adopts a dynamic equilibrium between an unliganded state, a dimeric complex state ((αII1)-GLOGEN), and a monomeric complex state (αII1-GLOGEN). Under conditions where the GLOGEN concentration is low with respect to the αII1, the dimeric complex state dominates, but as the peptide:αII1 ratio increases, the equilibrium position shifts to favor the formation of monomer. However, as collagen is abundant in the ECM and αI domains recognize heterotrimeric collagen sequences, we believe that this dimeric complex is unlikely to dominate in vivo and may not be biologically significant.

**CONCLUSION**

The structure of αII1 bound to GLOGEN suggests that the mechanism of signaling in this receptor is similar to that reported previously for α2I. However, the structure suggests that the mode of peptide binding to αII1 is subtly different from that observed with α2I. The current structure provides a rationale for the observed differences in binding specificity between αII1 and α2I.

---

**REFERENCES**

1. Hynes, R. O. (2002) Integrins: bidirectional, allosteric signaling machines. Cell 110, 673–687
2. Gullberg, D. E., and Lundgren-Akerlund, E. (2002) Collagen-binding 1 domain integrins—what do they do? Prog Histochem. Cytochem. 37, 3–54
3. Gullberg, D. E. (ed) (2003) I Domains in Integrins, pp. 25–57, Landes Bioscience, Georgetown, TX
4. Eble, J. A. (2005) Collagen-binding integrins as pharmaceutical targets. Curr. Pharm. Discov. 11, 867–880
5. Vanderslice, P., and Woodside, D. G. (2006) Integrin antagonists as therapeutics for inflammatory diseases. Expert Opin. Investig. Drugs 15, 1235–1255
6. Avraamides, C. J., Garmy-Susini, B., and Varner, J. A. (2008) Integrins in angiogenesis and lymphangiogenesis. Nat. Rev. Cancer 8, 604–617
7. Desgrosellier, J. S., and Cheresh, D. A. (2010) Integrins in cancer: biological implications and therapeutic opportunities. Nat. Rev. Cancer 10, 9–22
8. Michishita, M., Videm, V., and Arnaout, M. A. (1993) A novel divergent cation-binding site in the A domain of the β2 integrin CR3 (CD11b/CD18) is essential for ligand binding. Cell 72, 857–867
9. Tuckwell, D., Calderwood, D. A., Green, L. J., and Humphries, M. J. (1995) Integrin α2 I-domain is a binding site for collagens. J. Cell Sci. 108, 1629–1637
10. Qu, A., and Leahy, D. J. (1996) The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin. Structure 4, 931–942
11. Lee, J. O., Bankston, L. A., Arnaout, M. A., and Liddington, R. C. (1995) Two conformations of the integrin A-domain (I-domain): a pathway for activation? Structure 3, 1333–1340
12. Nykvist, P., Tu, H., Ivaska, J., Käpylä, I., Pihlajaniemi, T., and Heino, J. (2000) Distinct recognition of collagen subtypes by αβ1 and αβ1I integrins. αβ1 mediates cell adhesion to type XIII collagen. J. Biol. Chem. 275, 8255–8261
13. Käpylä, J., Ivaska, J., Rikkonen, P., Nykvist, P., Pentikäinen, O., Johnson, M., and Heino, J. (2000) Integrin α2I domain recognizes type I and type IV collagens by different mechanisms. J. Biol. Chem. 275, 3348–3354
14. Tulla, M., Pentikäinen, O. T., Viitasalo, T., Käpylä, J., Impola, U., Nykvist, P., Nissinen, L., Johnson, M. S., and Heino, J. (2001) Selective binding of collagen subtypes by integrin αII1, αII2, and αII10 domains. J. Biol. Chem. 276, 48206–48212
15. Knight, C. G., Morton, L. F., Peachey, A. R., Tuckwell, D. S., Farndale, R. W., and Barnes, M. J. (2000) The collagen-binding A-domains of integrins αβ1 and αβ1I recognize the same specific amino acid sequence, GFOGER, in native (triple-helical) collagens. J. Biol. Chem. 275, 35–40
16. Siljander, P. R., Hamaia, S., Peachey, A. R., Slater, D. A., Smethurst, P. A., Ouwehand, W. H., Knight, C. G., and Farndale, R. W. (2004) Integrin activation state determines selectivity for novel recognition sites in fibrillar collagens. J. Biol. Chem. 279, 47763–47772
17. Farndale, R. W., Lismann, T., Bihan, D., Hamaia, S., Smerling, C. S., Pugh, N., Konitsiotis, A., Leitinger, B., de Groot, P. G., Jarvis, G. E., and Raynal, N. (2008) Cell-collagen interactions: the use of peptide Toolkits to investigate collagen-receptor interactions. Biochem. Soc. Trans. 36, 241–250
18. Raynal, N., Hamaia, S. W., Siljander, P. R., Maddox, B., Peachey, A. R., Fernandez, R., Foley, L. J., Slater, D. A., Jarvis, G. E., and Farndale, R. W. (2006) Use of synthetic peptides to locate novel integrin α2β1-binding motifs in human collagen III. J. Biol. Chem. 281, 3821–3831
19. Xu, Y., Gurusiddappa, S., Rich, R. L., Owens, R. T., Keene, D. R., Mayne, R., Höök, A., and Höök, M. (2000) Multiple binding sites in collagen type I for the integrins αβ1 and αβ1I. J. Biol. Chem. 275, 38981–38989
20. Kim, J. K., Xu, Y., Xu, X., Keene, D. R., Gurusiddappa, S., Liang, X., Wary, K. K., and Höök, M. (2005) A novel binding site in collagen type III for integrins αβ1 and αβ1I. J. Biol. Chem. 280, 32512–32520
21. Hamaia, S. W., Pugh, N., Raynal, N., Némoz, B., Stone, R., Gullberg, D., Bihan, D., and Farndale, R. W. (2012) Mapping of potent and specific binding motifs, GLOGEN and GVOGEEA, for integrin αβ1 using collagen Toolkits II and III. J. Biol. Chem. 287, 26019–26028
22. Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J., and Liddington, R. C. (2000) Structural basis of collagen recognition by integrin αβ1. Cell
21. Gonzalez, J. L., Essafi, M., Xiong, J. P., Stehle, T., and Arnaout, M. A. (2002) Does the integrin αⅡb domain act as a ligand for its β3 domain? Curr. Biol. 12, R340–R342

22. Shimaoka, M., Xiao, T., Liu, J.-H., Yang, Y., Dong, Y., Jun, C.-D., McCormack, A., Zhang, R., Joachimiak, A., Takagi, J., Wang, J.-H., and Springer, T. A. (2003) Structures of the α1 domain and its complex with ICAM-1 reveal a shape-shifting pathway for integrin regulation. Cell 112, 99–111

23. Yang, W., Shimaoka, M., Salas, A., Takagi, J., and Springer, T. A. (2004) Intersubunit signal transmission in integrins by a receptor-like interaction with a pull spring. Proc. Natl. Acad. Sci. U.S.A. 101, 2906–2911

24. Connors, W. L., Jokinen, J., White, D. J., Puranen, J. S., Kankaanpää, P., Upla, P., Tulla, M., Johnson, M. S., and Heino, J. (2007) Two synergetic activation mechanisms of α2β1 integrin-mediated collagen binding. J. Biol. Chem. 282, 14675–14683

25. Aquilina, A., Korda, M., Bergelson, J. M., Humphries, M. J., Farndale, R. W., and Humphries, M. J., Farndale, R. W., and Barnes, M. J. (1998) Homology 2 domain studied by 15N NMR relaxation. Spectroscopy for real-time investigation of dynamic events in proteins on the microsecond time scale. J. Magn. Reson. Chem. 279, 7962–7970

26. Lee, J. H., and Woods, V. L. (2012) Dynamic structural changes are observed upon collagen and metal ion binding to the integrin α1 domain. J. Biol. Chem. 287, 32897–32912

27. Herrmann, T., Elsliger, M. A., Wilson, I. A., and Wuthrich, K. (2010) Comparison of NMR and crystal structures highlights conformational isomerism in protein active sites. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 66, 1393–1405

28. Herrmann, T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Wijg, J. S., Kuszewski, J. N., Gilmes, P. N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. 54, 905–921

29. Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J. Am. Chem. Soc. 125, 1731–1737

30. de Vries, S. J., van Dijk, A. D., Krzeminski, M., van Dijk, M., Thureau, A., Hsu, V., Wassenmaar, T., and Bonvin, A. M. (2007) HADDOCK versus HADDuck: new features and performance of HADDOCK2.0 on the CAPRI targets. Proteins 69, 736–733

31. Emsley, J., Knight, C. G., Farndale, R. W., and Barnes, M. J. (2004) Identification of the integrin α2β1-binding collagen peptide. J. Mol. Biol. 335, 1019–1028

32. Aquilina, A., Korda, M., Bergelson, J. M., Humphries, M. J., Farndale, R. W., and Humphries, M. J. (2002) Structures of the integrin α2β1 I domain carrying the activating mutation E317A. J. Biol. Chem. 286, 43343–43351

33. Shi, M., Pedchenko, V., Greer, B. H., Van Horn, W. D., Santoro, S. A., Sanders, C. R., Hudson, B. G., Eichman, B. F., Zent, R., and Pozzi, A. (2012) Enhancing integrin α2β1 inserted (I) domain affinity to ligand potentiates integrin α2β1-mediated down-regulation of collagen synthesis. J. Biol. Chem. 287, 35139–35152

34. Chiu, Y.-C., Headey, S., Mohnaty, B., Emley, J., Simpson, J. S., and Scallon, M. J. (January 22, 2013) Assignments of human integrin α1 domain in the apo and Mg2+ bound states. Biomol. NMR Assign. 10.1007/s11040-013-9465-7

35. Konarev, P. V., Petoukhov, M. V., Volkov, V. V., and Svergun, D. I. (2006) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Crystallogr. 39, 277–286

36. Lambert, L. J., Bobkov, A. A., Smith, J. W., and Marassi, F. M. (2008) Comparative interactions of collagen and a jararhagin-derived disintegrin peptide with the integrin α2β1 domain. J. Biol. Chem. 283, 16665–16672

37. Nymalm, Y., Puranen, J. S., Nyholm, T. K., Käpylä, J., Kidron, H., Pentikäinen, O. T., Airenne, T. T., Heino, J., Sottile, J. P., Johnson, M. S., and Smethurst, P. A., Tuckwell, D. S., Farndale, R. W., and Barnes, M. J. (1998) Identification in collagen type I of an integrin α2β1-binding site containing an essential GER sequence. J. Biol. Chem. 273, 33287–33294

38. Wishart, D. S., Rigam, C. G., Yao, J., Ablidgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) 1+H and 14N chemical shift referencing in biomolecular NMR. J. Biomol. NMR 6, 135–140

39. Goddard, T. D., and Kneller, D. G. (2001) SPARKY, 3rd Ed., University of California, San Francisco

40. Schanda, P., and Brutscher, B. (2005) Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. J. Am. Chem. Soc. 127, 8014–8015

41. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. J. Biomol. NMR 24, 171–189

42. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J. Mol. Biol. 319, 209–227

43. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. J. Mol. Biol. 273, 283–298

44. Serrano, P., Pedrini, B., Geralt, M., Jaudzems, K., Mohanty, B., Horst, R., Herrmann, T., Elsliger, M. A., Wilson, I. A., and Wüthrich, K. (2010) Comparison of NMR and crystal structures highlights conformational isomerism in protein active sites. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 66, 1393–1405

45. Domínguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J. Am. Chem. Soc. 125, 1731–1737

46. Herrmann, T., Guentert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. J. Biomol. NMR 24, 171–189

47. Herrmann, T., Guentert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J. Mol. Biol. 319, 209–227