



Cite this: *Org. Biomol. Chem.*, 2015, **13**, 8001

## Q2DSTD NMR deciphers epitope-mapping variability for peptide recognition of integrin $\alpha\nu\beta 6$ †

Jessica L. Sorge, Jane L. Wagstaff, Michelle L. Rowe, Richard A. Williamson\* and Mark J. Howard\*

Integrin  $\alpha\nu\beta 6$  is a cell surface arginine-glycine-aspartic acid (RGD)-specific heterodimeric glycoprotein that is only expressed on epithelia during processes of tissue remodelling, including cancer. The specificity and molecular nature of interactions toward this integrin are poorly understood and new insights into such processes are important to cell biologists and pharmaceutical drug discovery. This study demonstrates the application of quantitative two-dimensional saturation transfer (Q2DSTD) NMR to obtain precise details of peptide interactions with integrin  $\alpha\nu\beta 6$  and their correlation to specificity for the integrin. This approach highlights subtle but significant differences in ligand contact by three related 21-mer peptides: FMDV2, an  $\alpha\nu\beta 6$  specific peptide and DBD1 and LAP2T1 peptides that bind many  $\alpha\nu$  integrins in addition to  $\alpha\nu\beta 6$ . FMDV2 and DBD1 differ only by the cyclisation of DBD1; a process that removes  $\alpha\nu\beta 6$  specificity. Q2DSTD NMR demonstrates these peptides experience significantly different interactions with the integrin; FMDV contacts primarily through four residues: 6Leu, 10Leu, 12Val and 13Leu, whereas DBD1 and LAP2T1 have more widespread contacts across their sequences. Q2DSTD NMR combined two-dimensional STD with quantitation by considering the relaxation of the ligand (CRL) to provide precise ligand contact information. This study also examines the role of CRL in the Q2DSTD process and how quantitation modifies STD data and unravels epitope-mapping variability to provide precise results that differentiate interactions at the atomic level for each peptide.

Received 17th June 2015,  
Accepted 18th June 2015

DOI: 10.1039/c5ob01237f

www.rsc.org/obc

## Introduction

Saturation transfer difference (STD) NMR is ideal for studying the molecular detail of target-ligand recognition.<sup>1–3</sup> This study reports the use of Q2DSTD (Quantitative 2D saturation transfer difference); a combination of two-dimensional (2D) <sup>13</sup>C-edited STD NMR group epitope mapping<sup>2</sup> with added consideration of the relaxation of the ligand (GEM-CRL),<sup>4</sup> to distinguish subtle contact differences between three peptides with diverse specificities toward integrin  $\alpha\nu\beta 6$ . Group epitope mapping (GEM) is an approach where the relative intensities of STD signals provide information regarding the proximity of each ligand chemical group to the protein-target when it binds.<sup>2</sup> This approach is powerful because it informs on the closest and most significant molecular contacts that relate to ligand binding-mode and significant molecular interactions. Q2DSTD has the power to decipher subtle contact differences that are

difficult to distinguish using standard STD and GEM methods. 2D heteronuclear editing is required to unravel the large number of degenerate <sup>1</sup>H chemical shifts that are problematic from peptide samples containing several similar or identical amino acids. Considering the relaxation of the ligand involves subsequent data modification where <sup>1</sup>H longitudinal relaxation ( $T_1$ ) relaxation adjusts the STD contact result that can be distorted by structural anisotropy, non-specific aggregation and amphipathicity. The expected variability in <sup>1</sup>H  $T_1$  across each peptide sequence is influenced by amino acid composition and nascent structure and cannot be easily predicted. This study demonstrates advantages of  $T_1$  adjusted STD data for peptides to ascertain precise contact information to understand contacts responsible for specificity. It is most striking because we demonstrate QSTD contacts vary across three biologically distinct peptides that are all 21-mers of similar molecular weight. Only through analysis of precise quantitative STD data will significant contacts relevant to ligand specificity be identified.

Integrins are heterodimeric glycoproteins composed of non-covalently linked  $\alpha$  and  $\beta$  subunits<sup>5</sup> that dynamically translate extracellular matrix cues into intracellular responses

School of Biosciences, University of Kent, Canterbury, Kent, UK.

E-mail: m.j.howard@kent.ac.uk

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5ob01237f



to modulate cell proliferation, survival, migration and invasion.<sup>6</sup> Integrin  $\alpha\beta6$  is only expressed on epithelia during processes of tissue remodeling such as wound healing, inflammation and cancer.<sup>7,8</sup> Integrin  $\alpha\beta6$  has been identified as conveying a pro-invasive and aggressive phenotype when overexpressed on cancer cells;<sup>8–11</sup> it is an emerging clinical target as survival from cancers is reduced significantly if high levels of this integrin are expressed.<sup>12</sup> The real challenge revolves around finding ligands that bind only to  $\alpha\beta6$  that can be developed into medical agents and is complicated by  $\alpha\beta6$  being a arginine-glycine-aspartic acid (RGD) receptor class integrin that include  $\alpha\beta3$ ,  $\alpha\beta5$ ,  $\alpha\beta8$ ,  $\alpha8\beta1$ ,  $\alpha5\beta1$  and  $\alpha11\beta3$ ; all of which bind molecules that contain the RGD motif. STD NMR has previously been used to study integrin  $\alpha\beta6$ -peptide recognition.<sup>13–15</sup> 1D  $^1\text{H}$  and 2D  $^{13}\text{C}$ -edited STD NMR validated the affinity toward  $\alpha\beta6$  within peptides containing a turn-helix and a primary interacting extended RGDXXL/I motif. FMDV2 is the only RGD-peptide ligand known to bind exclusively to  $\alpha\beta6$ , but peptides DBD1 and LAP2T1 exhibit recognition to  $\alpha\beta6$  and a variety of other  $\alpha\text{v}$  integrins including  $\alpha\beta5$ ,  $\alpha\beta6$  and  $\alpha\beta8$ .<sup>16</sup>

The Q2DSTD approach was used to highlight molecular differences in peptide recognition that could be attributed to FMDV2's specificity for  $\alpha\beta6$ . What is equally interesting is the comparison between FMDV2 and DBD1; these peptides are related in sequence but structurally distinct through the disulfide-bond cyclisation of DBD1. The fact that FMDV2 is specific to  $\alpha\beta6$  but DBD1 is not supports the peptide–integrin interaction as extremely sensitive to both sequence and structure.<sup>16</sup> The three peptides used in this study were:

FMDV2  $\rightarrow$  NAVPNLRGDLQVLAQKVART-Hsl

DBD1  $\rightarrow$  EKCPNLRGDLQVLAQKVCRT-Hsl

LAP2T1  $\rightarrow$  GFTTGRRGDLATIHGLNRPF-Hsl

where Hsl is a C-terminal homoserine lactone produced as a result of recombinant expression and purification.<sup>17</sup> LAP2T1 is a modified latency associated peptide of TGF- $\beta1$  with a M16L mutation to enable recombinant production.<sup>16,17</sup>

## Results and discussion

$^{13}\text{C}$ -edited 2D STD NMR data, from samples containing identical concentrations of peptide and integrin, displayed the interaction of each peptide. The spectra labelled with key contact resonances shown in Fig. 1 are in agreement with those observed from previous studies<sup>13,14</sup> and peptide  $^1\text{H}$  and  $^{13}\text{C}$  assignments are tabulated in the ESI.† However, these previous studies provided no STD adjustment for ligand relaxation to provide quantitative contact analysis.

$^1\text{H}$   $T_1$  times measured using 2D  $^{13}\text{C}$ -edited inversion recovery varied between 0.29 and 0.64 s across all peptides (Fig. 2a) and displayed no correlation between sequences. The pulse sequence for 2D inversion recovery is shown in the ESI.† DBD1

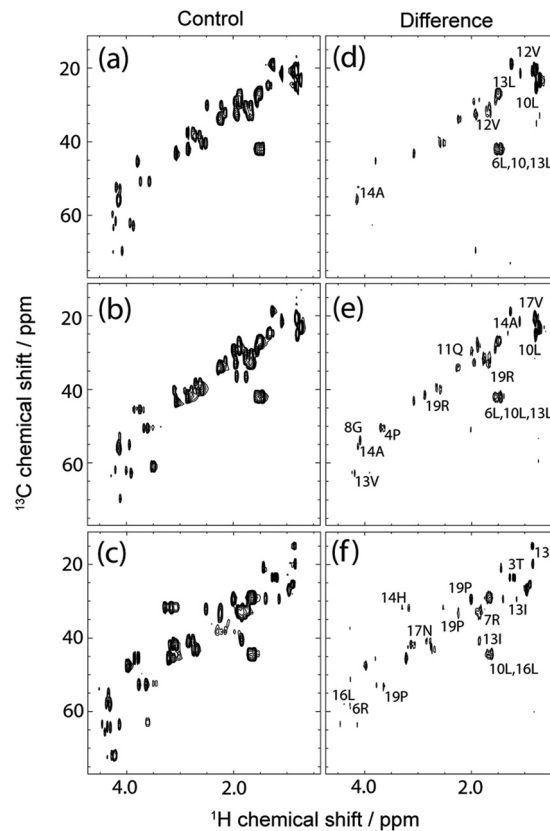
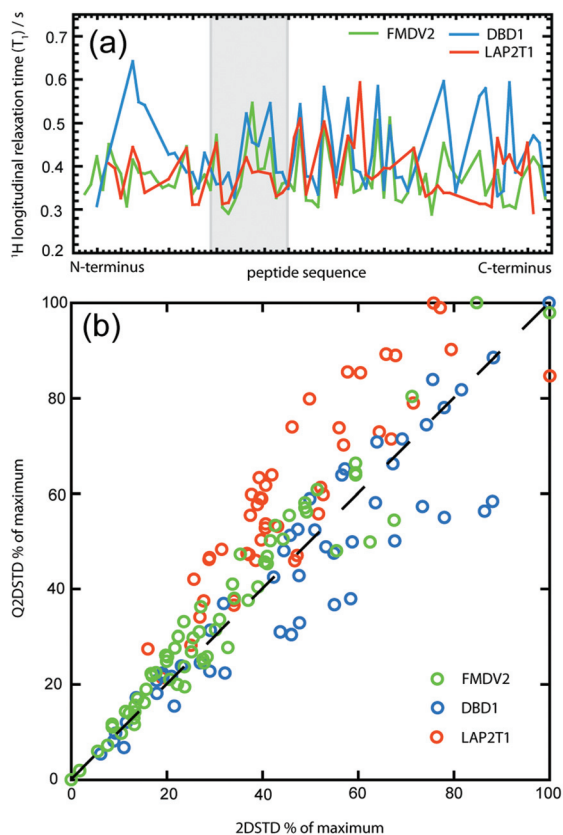


Fig. 1  $^{13}\text{C}$ ,  $^1\text{H}$ -STD HSQC spectra of the aliphatic regions for 1 mM ligands FMDV2 (a,d), DBD1 (b,e), and LAP2T1 (c,f) with 19  $\mu\text{M}$  integrin  $\alpha\beta6$ . Key resonances are labelled with residue number and data were acquired at 14.1 T and 10  $^\circ\text{C}$ .

provided the greatest variation in  $^1\text{H}$   $T_1$  across the sequence that most likely reflects the cyclic configuration of this peptide that contributes structural anisotropy and subsequent  $T_1$  variability. The influence of  $^1\text{H}$   $T_1$  on Q2DSTD is demonstrated in Fig. 2b, where the relative to maximum 2DSTD and Q2DSTD signals are plotted for all measured  $^1\text{H}$  in each peptide to demonstrate deviations caused by  $T_1$  relaxation. Proton longitudinal relaxation provides an average amendment of  $\pm 8.2\%$  from non-adjusted  $2\text{DSTD}_{\text{amp}}\%$  (standard deviation of 7.8%) but for the extremes  $2\text{DSTD}_{\text{amp}}\%$  was modified by up to  $\pm 30\%$ . Fig. 2b also illustrates the most significant differences occurring at higher  $2\text{DSTD}_{\text{amp}}\%$  values as well as LAP2T1 values being generally underestimated in contrast to DBD1 values that were overestimated. This supports the expected dependence of precision in STD being dependent on  $^1\text{H}$   $T_1$  values from the ligand. Also, the sequence similar but structurally distinct FMDV2 and DBD1 peptides produce very different Q2DSTD vs. STD correlations and confirms the need to handle peptide STD data with GEM modification where structural differences are known to exist. Fig. 2 clearly demonstrates that different peptides of the same length can exhibit enhanced or reduced Q2DSTD signals due to variability in  $^1\text{H}$   $T_1$  that subsequently influences each closest contact result at the atomic





**Fig. 2** The variability of  $^1\text{H}$  longitudinal relaxation time ( $T_1$ ) across all peptide sequences (a) with RGD motif data greyed out. Correlation graph highlighting differences in Q2DSTD% versus 2DSTD% for each measured data point in all peptides (b). FMDV2 data is shown in green, DBD1 data in blue and LAP2T1 data in red.

level.  $^1\text{H}$   $T_1$ , 2DSTD and Q2DSTD data for each peptide are fully tabulated in the ESI.†

However, detailed analysis of Q2DSTD in contrast to 2DSTD is shown in Fig. 3 where data from both approaches are shown across the RGDLLXL/I motif for each peptide. Arrows on the Q2DSTD data describe the result of  $T_1$  correction for Q2DSTD enhanced (†) or suppressed (↓). On a global level, the 2DSTD and Q2DSTD values are significantly different for each peptide, despite each experiment being acquired identically with the same concentrations. Global STD values for FMDV2 are largest with the smallest from DBD1 with LAP2T1 approximately an order of magnitude higher than values from the cyclic peptide. This is a reflection of the relative peptide binding affinities to integrin  $\alpha\beta 6$  that is in the order DBD1 > FMDV2 > LAP2T1<sup>16</sup> that is corroborated from integrin  $\alpha\beta 6$  IC<sub>50</sub> data for FMDV2 and LAP2T1 of 1.2  $\mu\text{M}$  and 13.8  $\mu\text{M}$  respectively.<sup>13</sup> Previous data<sup>13,16</sup> enabled an estimate of the IC<sub>50</sub> of DBD1 for  $\alpha\beta 6$  of ca. 0.5  $\mu\text{M}$ . STD NMR is extremely sensitive to  $K_d$  and associated off-rate<sup>3</sup> where the smaller the  $K_d$ , the smaller the off-rate observed. Off-rate is crucial for the detection of the ligand because it dictates the dissociation of the protein-ligand complex upon which saturation transferred

to the ligand whilst bound is finally observed. If the on-rate is diffusion limited at  $1 \times 10^{-7} \text{ s}^{-1} \text{ M}^{-1}$ , the off-rate is 0.1  $\text{s}^{-1}$ , 1.0  $\text{s}^{-1}$ , 10.0  $\text{s}^{-1}$  for  $K_d$  values of 10 nM, 100 nM and 1  $\mu\text{M}$  respectively. DBD1 has a small  $K_d$  and off-rate and produces weak STD values compared to FMDV2 that dissociates 10–100 times faster. STD data for LAP2T1 suggests this ligand may dissociate too quickly for optimal saturation transfer, hence its values are greater than DBD1 but less than FMDV2. As a result, it is likely that LAP2T1, with methionine at position 16 instead of leucine, is a weaker binder than the original LAP TGF- $\beta 1$   $\alpha\beta 6$  peptide.<sup>16</sup> These binding characteristics provide a STD and QSTD ‘sweet-spot’ that FMDV2 utilises to deliver the largest signals.

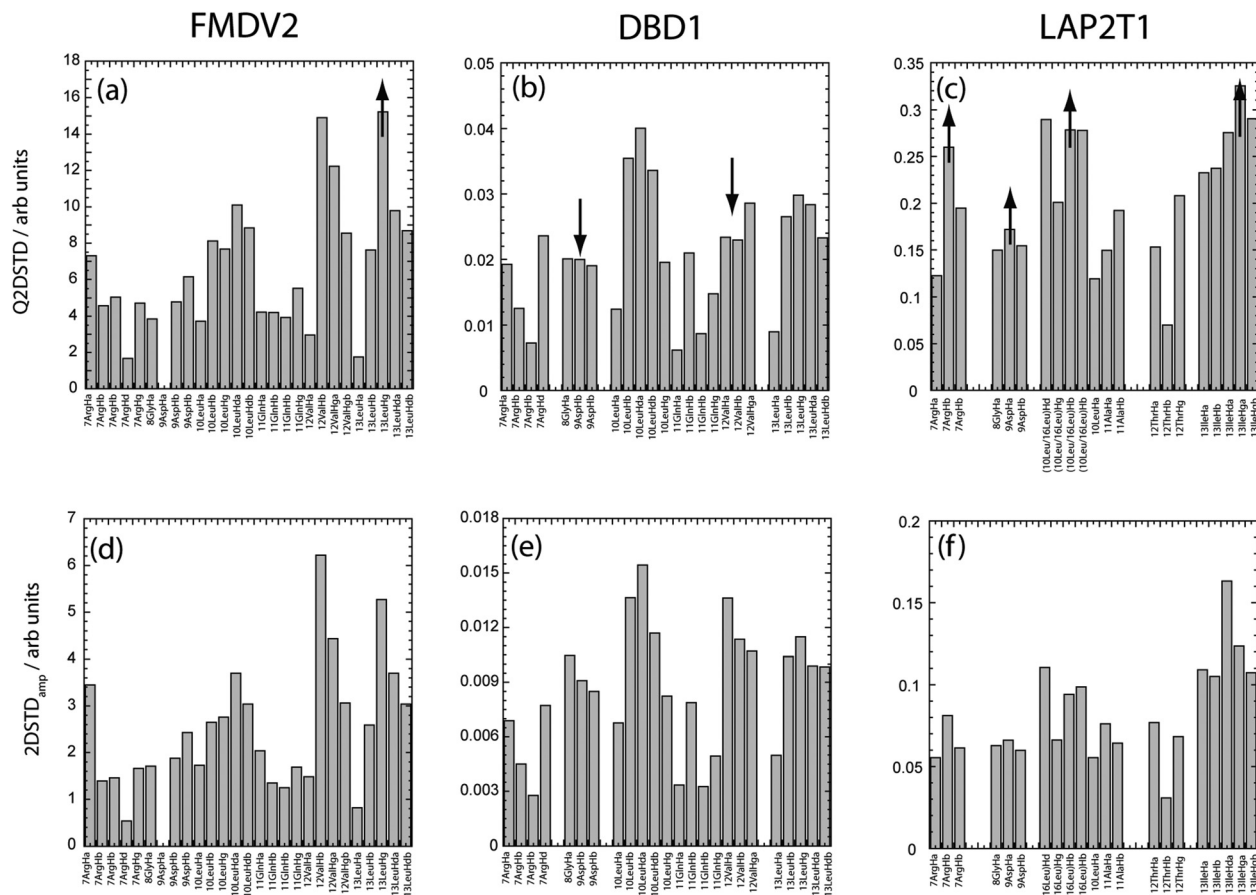
Fig. 3 confirms Q2DSTD reports different major contacts for each peptide; 13LeuH $\gamma$  for FMDV2, 10LeuH $\delta$  for DBD1 and 13IleH $\gamma$  for LAP2T1. The contact from 13LeuH $\gamma$  in FMDV2 provided the maximum Q2DSTD signal (Fig. 3a) and elevates the significance of this contact as greater than 12ValH $\beta$  that was superior in the non-adjusted 2DSTD (Fig. 3d). This Q2DSTD result is important because it is in contrast with previous studies that identify 12Val as the most significant contact residue. The elevation of 13Leu in FMDV2 is the only major change that occurs as a result of Q2DSTD analysis but both DBD1 and LAP2T1 provide additional data modifications that are of functional interest.

In contrast to FMDV2, Q2DSTD analysis of DBD1 data delivers maxima across 10Leu H $\beta$  and H $\delta$  protons and suppressed 2DSTD data points from 8Gly, 9Asp and 12Val when comparing Fig. 3b with Fig. 3e. The suppression of data upon quantitation is caused by relatively long  $^1\text{H}$   $T_1$  values for protons in these amino acids caused by the cyclic nature of this peptide. Suppression of 12Val signals in particular highlight the primary DBD1 contact with the integrin originating from 10Leu where Q2DSTD values ca. 1.3 times higher from 10Leu than 12Val or 13Leu. In contrast, FMDV2 produced Q2DSTD data that was ca. 1.6 times higher for 12Val and 13Leu than 10Leu. These observations enable differentiation of interactions between two similar peptides and integrin  $\alpha\beta 6$ . The disulphide bond cyclisation of DBD1 has influenced peptide primary contact with the integrin and changed specificity. FMDV2 is the only integrin  $\alpha\beta 6$  specific peptide published to date and it is also the only peptide to show significantly strong QSTD contacts toward the C-terminus of the RGDLLXL/I motif and weaker contacts from residues between positions 7–11.

LAP2T1 Q2DSTD data also produced significant modification of data between 2DSTD and Q2DSTD (Fig. 3c or f) with Q2DSTD showing enhancements across 7ArgH $\beta$ , 8GlyH $\alpha$ , 10/16LeuH $\beta$  and 13IleH $\gamma$ . Despite the ambiguous assignments for 10/16Leu, LAP2T1 clearly displays elevated 7Arg and 8Gly that confirms widespread major contacts across most of the RGDLLXL/I motif. Q2DSTD analysis alters the strongest amplification factor from 13IleH $\gamma$  to 13IleH $\delta$  in LAP2T1 that may sound insignificant, but not to pharmacophore design.

The combined Q2DSTD results from all three peptides separate FMDV2 as only interacting with integrin  $\alpha\beta 6$  through the C-terminus of the RGDLLXL/I motif; an observation coinci-





**Fig. 3** Comparison of calculated  $Q2DSTD_{amp}$  and  $2DSTD_{amp}$  values across the RGD/LXXL/I motif for FMDV2 (a,d), DBD1(b,e) and LAP2T1(c,f) following the interaction with integrin  $\alpha\beta6$ . Changes due to quantitation are depicted with the arrows in figures (a,b,c) and demonstrate how QSTD data has elevated ( $\uparrow$ ) or suppressed ( $\downarrow$ ) non quantitative STD results across individual protons in each peptide. FMDV2 and DBD1 have identical amino acid sequences across this motif region. Gaps are included to ensure residue positions are identically placed in each graph despite missing assignments, data or sequence differences (in the case of LAP2T1).

dent with its unique specificity for the integrin. Further analysis of Fig. 3a confirms this RGD section produces an average  $Q2DSTD$  of 4.3 compared to 15.2 for 13LeuH $\gamma$  in FMDV2. Therefore the RGD  $Q2DSTD$  average is only 28% of the maximum motif signal. In contrast, the average from the same RGD regions in DBD1 and LAP2T1 rise to 43% and 52% of the maximum motif signal respectively.

There are also advantages to viewing  $Q2DSTD$  across the entire sequence as shown in Fig. 4 with the results now reported for each peptide as percentage of their individual maximum signals. Amino acids for each peptide that contains at least one  $Q2DSTD\%$  value greater than the mean  $Q2DSTD\% + 0.25\sigma$  (50%) are shown as spheres to illustrate the extent of potentially significant contacts across each peptide. Using this approach Fig. 4 fully supports conclusions from Fig. 3 that FMDV2 has a smaller significant contact surface compared to DBD1 and LAP2T1. The 12Val-13Leu contact maxima from FMDV2 are distinct and easily identifiable and the ‘-KVART’ C-terminal region of this peptide produces  $Q2DSTD$  values

<50%. In contrast DBD1 and LAP2T1 show values >50% in their C-terminal residues and support the hypothesis that significant contact from peptides that are not  $\alpha\beta6$  specific is more widespread across their sequences. The same analysis can be made regarding the N-terminal regions of DBD1 and LAP2T1. The  $Q2DSTD$  data suggests the specific nature of FMDV2 toward  $\alpha\beta6$  is driven by a small number of specific contacts in the peptide, with only 6Leu, 10Leu, 12Val and 13Leu registering relative  $Q2DSTD\%$  values within 50% of maximum. Equivalent residues with  $Q2DSTD\% > 50\%$  are widespread across the sequences of DBD1 and LAP2T1.  $Q2DSTD\%$  data for the entire sequence of each peptide is shown in the ESI $\dagger$  and further supports a specific contact for FMDV2 and contacts over a larger proportion of both DBD1 and LAP2T1 peptides. In addition, previous STD studies noted that the RGD-motif registered weakly in FMDV2.<sup>13,14</sup> This was also observed in  $Q2DSTD$  data (Fig. 3 and ESI $\dagger$ ) and confirms the original observation was not as a manifestation of  $^1H T_1$  effects. FMDV2 also exhibits a uniquely weak 8GlyH $\alpha$  inter-



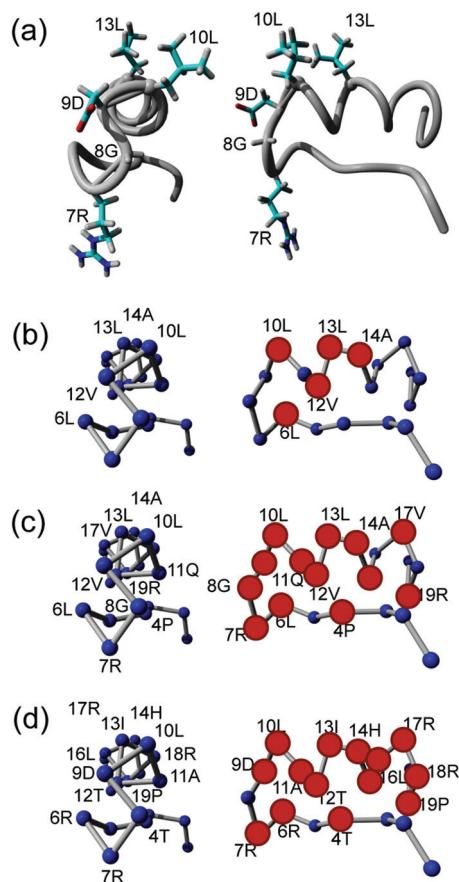


Fig. 4 Backbone view of the FMDV2 structure with the RGDLLXL motif highlighted (a).<sup>13,16</sup> Principal Q2DSTD contacts ( $>Q2DSTD\% + 0.25\sigma$ ) are labelled for FMDV2 (b), DBD1 (c) and LAP2T1 (d) and highlighted as large red spheres in the side on structure. This supports the wider contact by DBD1(c) and LAP2 (d) and specific contact by FMDV2 (b).

action that could also explain its poor specificity for other RGD-specific  $\alpha$ v integrins. This hypothesis was alluded to in earlier 1D STD data<sup>13</sup> but was previously difficult to confirm due to spectral overlap.

Using QSTD2D to compare the extended FMDV2 peptide to the disulphide bond cyclized DBD1 highlights similar patterns of interaction. However, the DBD1 peptide has a prominent ( $>50\%$  QSTD<sub>amp</sub>) RGD interaction that is not reflected in the FMDV2 data and could explain the promiscuity of DBD1 for other  $\alpha$ v integrins. This is most likely the result of structural or conformational changes induced by cyclisation and supports the previous observation linking the rigidity of FMDV2 to  $\alpha$ v $\beta$ 6 specificity as well as conformational exchange to the loss of specificity by DBD1.<sup>16</sup> Furthermore, and whatever the effects of cyclic formation, DBD1 has functional behaviour akin to LAP2T1 and LAP2T1 was also shown to be conformationally active.<sup>16</sup>

Close examination of individual contacts across the RGDLLXL/I motif (Fig. 3) highlight the C-terminal motif preference for primary contact to the integrin by FMDV2. This detail informs that integrin  $\alpha$ v $\beta$ 6 specificity involves a small

binding site on the integrin and that FMDV2 is capable of accessing this binding site and also suppressing its interactions with other  $\alpha$ v integrins. This does suggest that shortening the FMDV2 sequence would create smaller peptides with equivalent properties but preliminary data suggests both  $\alpha$ v $\beta$ 6 specificity and affinity are lost upon FMDV2 truncation at either N- or C-terminus. This supports the original hypothesis that specific peptides for integrin  $\alpha$ v $\beta$ 6 require a turn-helix motif with a stabilised helical region beyond the RGDLLXL/I motif.<sup>13</sup>

## Conclusions

Q2DSTD is a powerful approach using isotopically enriched peptides to provide resolution, combined with quantitative correction of ligand longitudinal relaxation. This study demonstrated a valuable application of Q2DSTD NMR in providing specific contact information of similar peptides that highlight molecular specificity and recognition toward integrin  $\alpha$ v $\beta$ 6. These specific contacts from the FMDV2 peptide provide the first insight into the design of pharmacophores to create future pharmaceutical agents against this cancer integrin. In this study, the molecular features required for integrin  $\alpha$ v $\beta$ 6 specificity can be seen in Fig. 4b and need to mirror contacts made by 6Leu, 10Leu, 12Val, 13Leu and 14Ala with the primary contact point based in the vicinity of 13Leu.

## Experimental

### Recombinant peptide production

All laboratory reagents reagent grade or higher and supplied by Sigma-Aldrich unless otherwise stated. The production and purification of recombinant peptides is described below and also in our previous publications.<sup>16,17</sup> FMDV2 originates from the surface GH loop motif from foot and mouth disease virus serotype O<sub>1</sub> BFS capsid protein VP1.<sup>18</sup> DBD1 is a variant of FMDV2 with a cyclizing disulfide bond to limit serum degradation for clinical *in vivo* use. LAP2T1 is a natural  $\alpha$ v $\beta$ 6 ligand sequence from the latency-associated peptide (LAP) of transforming growth factor- $\beta$ 1 (TGF $\beta$ 1).<sup>13</sup>

The recombinant peptide production process uses the pET-31b vector in *E. coli* that utilises ketosteroid isomerase as the fusion partner with the peptide to facilitate easy extraction with minimal loss of the peptide due to proteolysis. This is achieved because the expression product is insoluble and forms inclusion bodies of the product. These inclusion bodies are easily purified, then solubilised and cleaved to yield peptide that is finally purified to  $>90\%$  by HPLC. Isotopic enrichment involves the growth of *E. coli*, complete with the pET-31b plasmid encoding the peptide, in a minimal medium with uniform labelled <sup>13</sup>C-glucose; the same approach used for making isotopically enriched proteins.

Sense and anti-sense oligonucleotides (MWG) were designed to encode peptide sequences for FMDV2, DBD1 and



LAP2T1. Oligonucleotides were 5' phosphorylated and designed with additional 3' overhangs of ATG for the sense and CAT for the anti-sense sequence so that annealed DNA could be ligated into the pET31b(+) (Novagen) vector pre-cut with *Aha*NI restriction enzyme. Oligonucleotides were inserted downstream of the N-terminal fusion protein ketosteroid isomerase and upstream of a C-terminal His-tag. Ligation mixtures were transformed into competent *E. coli* DH5 $\alpha$  cells and selected for by plating on ampicillin LB agar plates. Colonies were screened for oligonucleotide insertion by PCR or restriction digestion of purified plasmid using Xba I and Xho I restriction enzymes. Sequences of plasmids containing multiple inserts were sequenced and if correct transformed into competent *E. coli* BL21(DE3) cells ready for recombinant protein expression.

Recombinant  $^{13}\text{C}$  isotopically enriched fusion protein was expressed in minimal M9 medium at 37 °C at 200 rpm with  $^{15}\text{N}$  ammonium sulphate (Cambridge Isotopes, USA) as the sole nitrogen source. Protein expression was induced by addition of IPTG to a final concentration of 1 mM for 3–4 h when the OD<sub>600</sub> nm of the culture was between 0.55 and 0.7. Cells were harvested by centrifugation (15 min, 6300g) and the cell pellet re-suspended in lysis buffer (20 mM NaH<sub>2</sub>PO<sub>4</sub>, 50 mM NaCl, pH 7.3; 10 mL per 400 mL of original culture volume) and frozen. After thawing, cell lysis was completed by the addition of lysozyme to a final concentration of 0.01 mg mL<sup>-1</sup> and Triton X-100 at 0.1% v/v and incubated at RT for 20 min followed by the addition of 0.02 mg mL<sup>-1</sup> DNase I and 10 mM MgCl<sub>2</sub> until the viscosity of the solution was reduced followed by 2 min of pulsed sonication on ice. Insoluble fusion protein was then recovered from the total cell lysate by centrifugation (10 min, 12 000g) and purified by re-suspension in wash buffer (50 mM Tris-HCl, 10 mM EDTA, 0.5% Triton X-100, pH 8; 2.5 mL per 400 mL of original culture volume) and recovering by centrifugation (10 min, 12 000g), with this step repeated again with wash buffer and then a further two times with dH<sub>2</sub>O.

Purified peptide-protein inclusion bodies were solubilised with 6 mL of 85% formic acid and peptide released from the fusion by addition of 0.2 g cyanogen bromide, incubated in the dark at RT for 16–24 h. After incubation the formic acid solution was diluted with 20 mL of dH<sub>2</sub>O and lyophilised. Soluble peptides were then separated from the insoluble KSI stirring overnight in PBS (25 mM Na<sub>2</sub>HPO<sub>4</sub>, 100 mM NaCl; 2.5 mL per 400 mL original culture volume) the pH corrected to 7.5 and recovered by centrifugation. Peptide was then separated and purified from the contaminating His-tag using a Waters 600/486 series HPLC with a preparative Vidac C18 reverse phase protein and peptide column using an elution gradient of HPLC grade water and 70% acetonitrile/30% water containing 0.05% and 0.045% trifluoroacetic acid (TFA) respectively. Peptide containing fractions were collected when the absorbance of the flow at 220 nm reached 0.1 AU and stopped when the absorbance returned to 0.2 AU to minimise the risk of sample contamination. Collected fractions were then lyophilised to recover the peptide.

## NMR spectroscopy

All data were acquired at 10 °C using a 4-channel Varian UnityINOVA NMR spectrometer operating at 14.1 T (600 MHz  $^1\text{H}$ ) with a 5 mm room temperature HCN probe. 2D  $^{13}\text{C}$ ,  $^1\text{H}$ -STD-HSQC experiments used a previously published pulse sequence,<sup>14</sup> with Gaussian-based saturation<sup>19</sup> at -3 ppm and -30 ppm. These experiments were acquired with 2048 data points (8000 Hz) in the direct F2 dimension and 64 data pairs - 128 points (18 000 Hz) in the indirect F1 dimension.  $^{13}\text{C}$ ,  $^1\text{H}$ -HSQC spectra for assignment were acquired with 512 complex pairs (1024 points) in F1 and each inversion recovery  $^{13}\text{C}$ ,  $^1\text{H}$ -HSQC dataset were acquired with 128 complex pairs (256 points) in F1; all with  $^{13}\text{C}$  spectral width of 18 000 Hz. In addition, inversion recovery experiments utilised a scan recycle delay of 5 seconds.

## STD and QSTD NMR analysis

Quantitative STD analysis utilises  $^1\text{H}$  longitudinal relaxation times ( $T_1$ ) of protons in each CH<sub>*n*</sub> correlation to quantify 2D STD data. A comparison of Q2DSTD and 2DSTD data for all three peptides is shown in Fig. 1 and data for each resonance tabulated in ESI† observed.  $^1\text{H}$   $T_1$  relaxation times were acquired using a modified  $^{13}\text{C}$ ,  $^1\text{H}$ -HSQC experiment with a  $^1\text{H}$  inversion-recovery segment (180°- $\tau$ -90°) in place of the initial 90°  $^1\text{H}$ -pulse as shown in the ESI.† The inter-sequence relaxation delay was set to 5 s and  $\tau$  values were 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 1.0 s to provide intensities for each  $^1\text{H}$  resonance correlated to its  $^{13}\text{C}$  in the modified HSQC. Measurement of each series of intensities ( $I$ ) for each  $^1\text{H}$  resonance enabled the determination of  $^1\text{H}$   $T_1$  by curve-fitting the equation  $I = I_0[1 - 2e(-\tau/T_1)]$ . Two-dimensional quantitative STD (Q2DSTD) for each  $^1\text{H}$  were obtained using the equations below with intensities from control (2DSTD<sub>ctrl</sub>) and difference (2DSTD<sub>diff</sub>) experiments<sup>14</sup> where the ligand excess was 53-fold.

$$2\text{DSTD}_{\text{amp}} = [2\text{DSTD}_{\text{diff}}/2\text{DSTD}_{\text{ctrl}}] \times \text{ligand excess}$$

$$\text{Q2DSTD} = 2\text{DSTD}_{\text{amp}}/T_1$$

## Acknowledgements

MJH and RAW would like to thank the BBSRC for a studentship to JLW.

## Notes and references

- M. Mayer and B. Meyer, *Angew. Chem., Int. Ed.*, 1999, **38**, 1784–1788.
- M. Mayer and B. Meyer, *J. Am. Chem. Soc.*, 2001, **123**, 6108–6117.
- B. Meyer and T. Peters, *Angew. Chem., Int. Ed.*, 2003, **42**, 864–890.
- S. Kemper, M. K. Patel, J. C. Errey, B. G. Davis, J. A. Jones and T. D. Claridge, *J. Magn. Reson.*, 2010, **203**, 1–10.
- R. O. Hynes, *Cell*, 2002, **110**, 673–687.



- 6 C. A. Morton, I. Campbell and R. M. MacKie, *Br. J. Dermatol.*, 1996, **135**, 853–855.
- 7 D. Kohda, C. J. Morton, A. A. Parkar, H. Hatanaka, F. M. Inagaki, I. D. Campbell and A. J. Day, *Cell*, 1996, **86**, 767–775.
- 8 C. J. Morton, D. J. Pugh, E. L. Brown, J. D. Kahmann, D. A. Renzoni and I. D. Campbell, *Structure*, 1996, **4**, 705–714.
- 9 G. J. Thomas, M. P. Lewis, I. R. Hart, J. F. Marshall and P. M. Speight, *Int. J. Cancer*, 2001, **92**, 641–650.
- 10 G. J. Thomas, M. P. Lewis, S. A. Whawell, A. Russell, D. Sheppard, I. R. Hart, P. M. Speight and J. F. Marshall, *J. Invest. Dermatol.*, 2001, **117**, 67–73.
- 11 G. J. Thomas, M. L. Nystrom and J. F. Marshall, *J. Oral. Pathol. Med.*, 2006, **35**, 1–10.
- 12 B. Dalhus, M. Saarinen, U. H. Sauer, P. Eklund, K. Johansson, A. Karlsson, S. Ramaswamy, A. Bjork, B. Synstad, K. Naterstad, R. Sirevag and H. Eklund, *J. Mol. Biol.*, 2002, **318**, 707–721.
- 13 D. Dicara, C. Rapisarda, J. L. Sutcliffe, S. M. Violette, P. H. Weinreb, I. R. Hart, M. J. Howard and J. F. Marshall, *J. Biol. Chem.*, 2007, **282**, 9657–9665.
- 14 J. L. Wagstaff, S. Vallath, J. F. Marshall, R. A. Williamson and M. J. Howard, *Chem. Commun.*, 2010, **46**, 7533–7535.
- 15 B. Claasen, M. Axmann, R. Meinecke and B. Meyer, *J. Am. Chem. Soc.*, 2005, **127**, 916–919.
- 16 J. L. Wagstaff, M. L. Rowe, S.-J. Hsieh, D. DiCara, J. F. Marshall, R. A. Williamson and M. J. Howard, *Rsc. Adv.*, 2012, **2**, 11019–11028.
- 17 J. L. Wagstaff, M. J. Howard and R. A. Williamson, *Mol. Biosyst.*, 2010, **6**, 2380–2385.
- 18 D. Logan, R. Abu-Ghazaleh, W. Blakemore, S. Curry, T. Jackson, A. King, S. Lea, R. Lewis, J. Newman, N. Parry, *et al.*, *Nature*, 1993, **362**, 566–568.
- 19 N. B. Ley, M. L. Rowe, R. A. Williamson and M. J. Howard, *RSC Adv.*, 2014, **4**, 7347–7351.

