

This document is confidential and is proprietary to the American Chemical Society and its authors. Do not copy or disclose without written permission. If you have received this item in error, notify the sender and delete all copies.

### Molecular Simulation of $\alpha V\beta 6$ Integrin Inhibitors

Journal:	<i>Journal of Chemical Information and Modeling</i>
Manuscript ID	ci-2020-00254f.R4
Manuscript Type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Guest, Ellen; University of Nottingham, School of Chemistry Oatley, Steven; University of Nottingham, School of Chemistry Macdonald, Simon; GlaxoSmithKline Medicines Research Centre, Medicinal Science & Technology Hirst, Jonathan; University of Nottingham, School of Chemistry

SCHOLARONE™  
Manuscripts

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

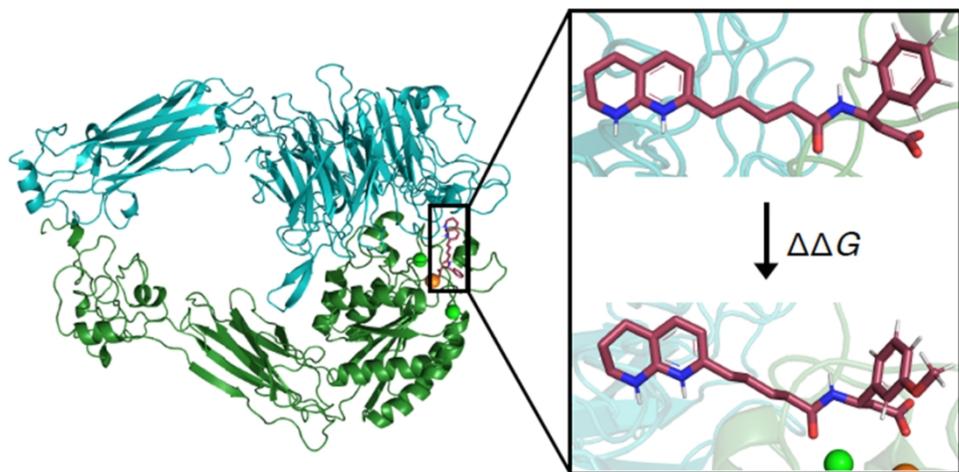


Table of Contents graphic

338x190mm (96 x 96 DPI)

# Molecular Simulation of $\alpha v \beta 6$ Integrin Inhibitors

Ellen E. Guest,<sup>†</sup> Steven A. Oatley,<sup>†</sup> Simon J. F. Macdonald,<sup>‡</sup> and Jonathan D.

Hirst<sup>\*,†</sup>

<sup>†</sup>*School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, U.K.*

<sup>‡</sup>*RGDS Science Ltd, 2 Walsworth Road, Hitchin, Herts SG4 9SP*

E-mail: jonathan.hirst@nottingham.ac.uk

## Abstract

The urgent need for new treatments for the chronic lung disease idiopathic pulmonary fibrosis (IPF) motivates research into antagonists of the RGD binding integrin  $\alpha v \beta 6$ , a protein linked to the initiation and progression of the disease. Molecular dynamics (MD) simulations of  $\alpha v \beta 6$  in complex with its natural ligand, pro-TGF- $\beta 1$ , show the persistence over time of a bidentate Arg-Asp ligand-receptor interaction and a metal chelate interaction between an aspartate on the ligand and an  $Mg^{2+}$  ion in the active site. This is typical of RGD binding ligands. Additional binding site interactions, which are not observed in the static crystal structure, are also identified. We investigate an RGD mimetic, which serves as a framework for a series of potential  $\alpha v \beta 6$  antagonists. The scaffold includes a derivative of the widely utilised 1,8-naphthyridine moiety, for which we present force field parameters, to enable MD and relative free energy perturbation (FEP) simulations. The MD simulations highlight the importance of hydrogen bonding and cation- $\pi$  interactions. The FEP calculations predict relative binding affinities, within  $1.5 \text{ kcal mol}^{-1}$ , on average, of experiment.

## Introduction

Integrin  $\alpha v \beta 6$  is a transmembrane, heterodimeric protein. The involvement of integrins in the progression of several tumour types and diseases including idiopathic pulmonary fibrosis (IPF) makes them important targets for the development of new drug compounds.<sup>1-3</sup> IPF is a chronic disease characterised by progressive scarring of the lungs. There are no curative treatments for IPF and survival rates are two to four years upon diagnosis.<sup>4</sup> The involvement of  $\alpha v \beta 6$  in the initiation and progression of IPF motivates research to understand the conformational changes and binding properties of  $\alpha v \beta 6$  with its natural ligand, the pro-domain of TGF- $\beta 1$ , and a series of potential antagonists.<sup>5</sup> Latent TGF- $\beta 1$  is activated upon binding to integrin  $\alpha v \beta 6$  through an Arg-Gly-Asp (RGD)LXX(I/L) motif (where X is any amino acid) in the pro-domain. Activated TGF- $\beta 1$  responds to tissue injury, infection and mediates tissue repair. However, increased activity leads to tissue inflammation and scar formation in uninjured areas.<sup>6,7</sup> A drug antagonist of  $\alpha v \beta 6$  could treat IPF through active site binding in place of TGF- $\beta 1$ .

Identifying new clinical candidates which balance all necessary properties and that can be administered as a low dose medicine, is difficult. An important aspect, particularly for the clinical dose size, is the affinity of the modulator for its biological target. The more potent the molecule is whilst controlling its lipophilicity, the greater the chance of a lower clinical dose. This study lays the foundation for computationally estimating the affinity of an inhibitor for the  $\alpha v \beta 6$  integrin. Predicting ligand affinity in integrin drug discovery from docking studies has historically been difficult (although it is generally possible to rationalise the activity once the data is available<sup>8</sup>). Therefore, understanding which structural features of the inhibitor are most important for driving affinity from molecular simulations complements the empirical process that is generally used in integrin lead optimisation where compounds are made, tested and structure activity relationships (SAR) are developed. Computational predictions become particularly important with longer syntheses, which use considerable resources and take months to complete.

The  $\alpha v$  subunit of extracellular  $\alpha v \beta 6$  comprises a lower and upper calf domain, the thigh domain and the  $\beta$ -propeller domain (Figure 1). Extracellular  $\beta 6$  is made up of a lower leg, upper

leg, and head region. The head region consists of the plexin-semaphorin-integrin (PSI) domain, the hybrid domain and the  $\beta 1$  domain. Ligand binding occurs at the interface of the  $\alpha v$  and  $\beta 6$  subunit, with interactions formed between the ligand and each of the head regions (Figure 1). The activity of  $\alpha v \beta 6$  is influenced by conformational changes in the multidomain subunits of the integrin. In its inactive bent-closed conformation, the subunits hinge at each Genu, so that the head regions fold towards the cell membrane. Activation involves a switchblade-like opening motion of the upper domains to give the extended-closed conformation, exposing the active site to ligand binding. A swing out motion of the hybrid domain away from the  $\alpha v$  subunit induces the extended-open conformation. This results in a more closed binding site at the other end of the  $\beta 1$  domain.

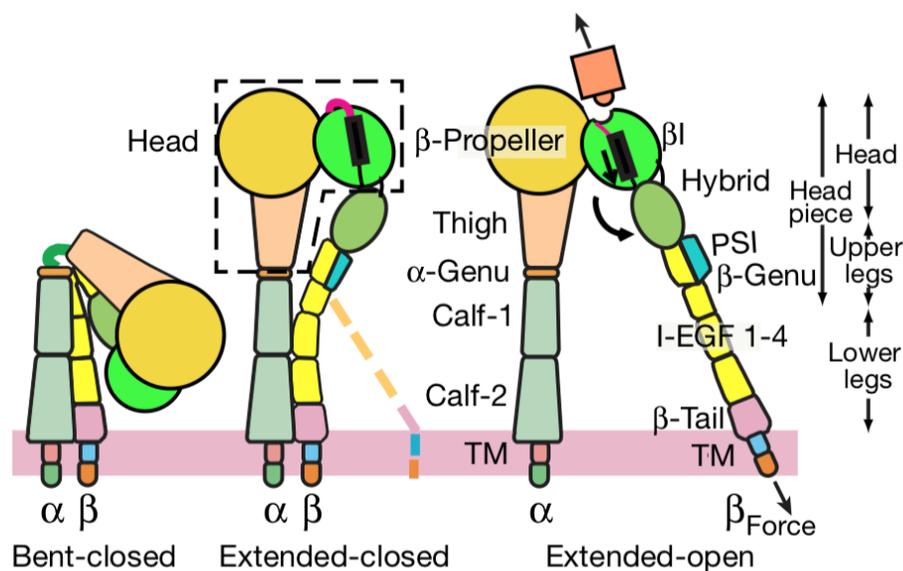


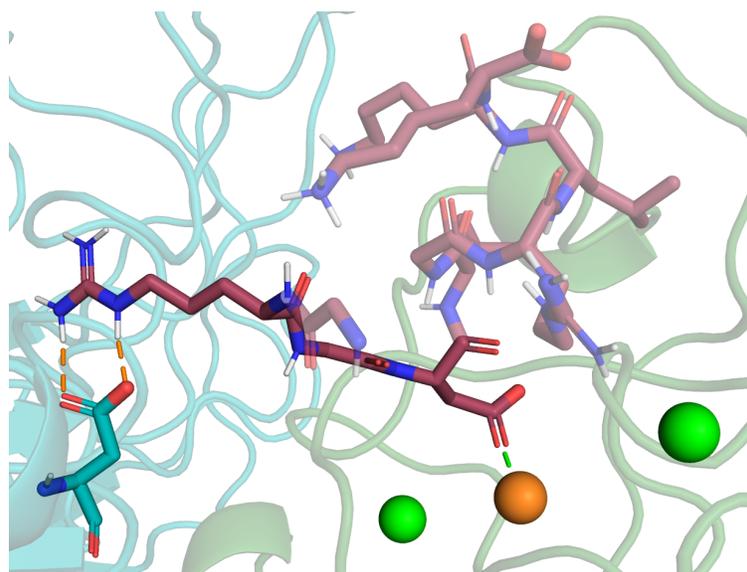
Figure 1: Schematic showing the different conformations of  $\alpha v \beta 6$ . The domains of the integrin are labelled with the extracellular protein shown above the membrane (denoted as a pink strip); the ligand is shown in orange.<sup>9</sup> [Reproduced with permission from ref 9.]

Pro-TGF- $\beta 1$  binds to  $\alpha v \beta 6$  through an RGD motif. The Arg<sup>RGD</sup> side chain interacts with the carboxyl group on ( $\alpha v$ )-Asp<sup>218</sup> through bidentate hydrogen bonds (Figure 2a). The carboxyl group on Asp<sup>RGD</sup> coordinates with a Mg<sup>2+</sup> ion in the metal ion-dependent adhesion site (MIDAS) of the  $\beta 6$  subunit. As these interactions are consistent across all  $\alpha v$  integrins bound to RGD ligands, we term this set of interactions as canonical interactions. The Mg<sup>2+</sup> ion is flanked by two Ca<sup>2+</sup>

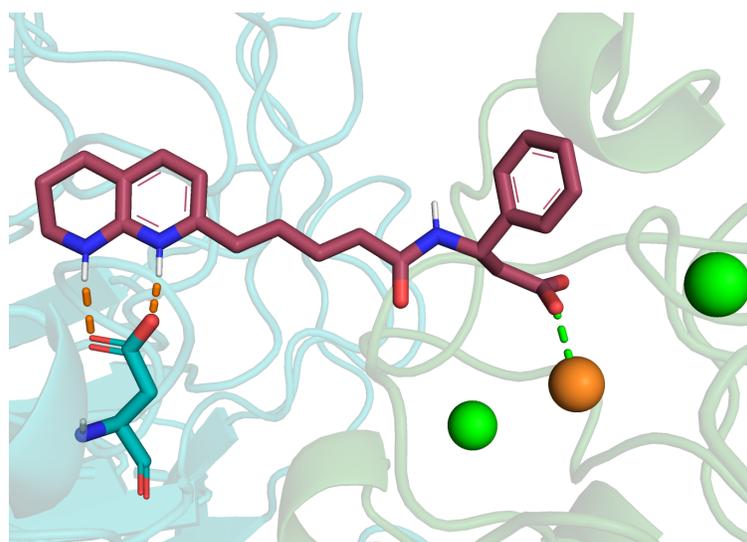
ions, one of which is called the adjacent to MIDAS (ADMIDAS). The MIDAS and ADMIDAS are important contributors to ligand binding and conformational rearrangements of the receptor. The opening of the hybrid domain to give the extended-open conformation results in a 3 Å movement of the ADMIDAS towards the MIDAS.<sup>9</sup> This concerted movement decreases the accessibility of the binding site in the extended-open conformation. The X-ray crystal structure with Protein Data Bank (PDB) code 4UM9<sup>10</sup> has been selected for our study. In this structure,  $\alpha\beta6$  is in complex with TGF- $\beta$ 1 and has a complete MIDAS and ADMIDAS occupancy. In addition to the canonical interactions displayed in the crystal structure, there are other binding site contacts. The RGD aspartate on the ligand forms hydrogen bonds with ( $\beta$ 6)-Asn<sup>218</sup> and ( $\beta$ 6)-Ala<sup>126</sup> in the receptor. The backbone of the ligand and ( $\beta$ 6)-Thr<sup>221</sup> also interact.

By taking advantage of non-canonical intramolecular interactions determined through molecular dynamics (MD) simulations, Di Leva et al. identified a  $\alpha\beta6$  potent cyclic peptide from an RGD-containing linear oligomer.<sup>12</sup> This illustrates the utility of MD simulations for identifying potential areas for ligand development through a detailed description of how the dynamic changes of active site residues contribute to receptor-ligand binding. A computational study of  $\alpha\beta3$  - lig<sup>RGD</sup> complexes found the Arg<sup>RGD</sup> - ( $\alpha$ v)-Asp<sup>218</sup> interaction is maintained over 100 ns.<sup>13</sup> Over the simulation the distance between Asp<sup>RGD</sup> and the Mg<sup>2+</sup> ion in the MIDAS of  $\beta$ 3 decreased, indicating these interactions are stable. Another MD simulation study<sup>14</sup> revealed, by varying isoDGR-containing cyclopeptides in complex with  $\alpha\beta3$ , subtle differences in  $\alpha\beta3$ -ligand interactions that affect the allosteric response of the receptor to ligand binding.

Our interest focuses on an RGD mimetic as the framework for a series of potential antagonists, as shown in Figure 2b. A 1,2,3,4-tetrahydro 1,8-naphthyridine group at one end of the compound mimics the Arg residue in the RGD tripeptide. This moiety is of particular interest in medicinal chemistry as 1,8-naphthyridines and their derivatives are found in many natural substances with biological activities.<sup>15</sup> The carboxyl on the opposite end of the compound mimics the Asp residue and binds to the Mg<sup>2+</sup> MIDAS ion. Each compound varies at the 3-position of the aryl group, with substituents including H, F, CF<sub>3</sub>, OCF<sub>3</sub>, CH<sub>3</sub> and OCH<sub>3</sub>. These compounds were taken from



(a)



(b)

Figure 2: TGF- $\beta$ 1 (a) and the RGD mimetic (b) bound in the active site of  $\alpha$ v $\beta$ 6, showing the  $\alpha$ v subunit (teal), the  $\beta$ 6 subunit (faded green), the backbone of the ligands (maroon) and heteroatoms (blue and red). The RGD sequence of TGF- $\beta$ 1 is shown in bold. The bidentate ( $\alpha$ v)-Asp<sup>218</sup> hydrogen bonding and the metal chelate interaction with the Mg<sup>2+</sup> ion (orange) are indicated by dashed lines. The Ca<sup>2+</sup> ions are shown in green. Image generated with the PyMol Molecular Graphics System.<sup>11</sup>

1  
2  
3 a class of compounds in an ongoing study of the SAR of  $\alpha_v$  integrin antagonists.<sup>5,16</sup> Although  
4 other chemotypes have been explored in literature,<sup>8,17</sup> our study focuses on a single scaffold.  
5 Scaffold substituents were chosen in our study as they are compatible with the CHARMM force  
6 field, obviating the need for extensive force field development for individual compounds. In  
7 previous work, the activity of each compound has been measured through a cell adhesion assay.  
8 The compounds have pIC<sub>50</sub> values in the range of 5.2 to 7.1. Due to the racemic nature of the  
9 compound, a single pIC<sub>50</sub> value is assigned to both the (*R*)- and (*S*)-enantiomer of each compound,  
10 with the exception of CF<sub>3</sub>. As the CF<sub>3</sub> analogue was prepared from commercially available (*R*)  
11 and (*S*) precursors, distinct pIC<sub>50</sub> values can be assigned to each enantiomer. pIC<sub>50</sub> is the negative  
12 log<sub>10</sub> of the half maximal inhibitory concentration (IC<sub>50</sub>) of a compound. We can relate IC<sub>50</sub> to  
13  $\Delta G$  using Equation 1.  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24

$$\Delta G = -RT \ln(\text{IC}_{50}) \quad (1)$$

25  
26  
27  
28 In this study, we build on previous calculations<sup>18</sup> using molecular docking. Starting from  
29 docked conformations, we use MD and free energy perturbation (FEP) simulations<sup>19</sup> to assess the  
30 effects of the different aryl substituents on active site interactions. There is growing interest in  
31 predicting relative binding energies using FEP simulations and integrating them into drug discovery  
32 workflows.<sup>20</sup> Chemically accurate *in silico* binding affinities can provide guidance when optimising  
33 lead compounds, saving synthetic resources and effort. There is a question, however, about the  
34 ease with which FEP can be applied to new systems. Nevertheless, with improved force fields  
35 and greater computational resources becoming more available, FEP simulations are increasingly  
36 attractive and tractable.  
37  
38  
39  
40  
41  
42  
43  
44  
45

46 The accuracy of FEP simulations rely on a quality force field, sufficient sampling and a well  
47 equilibrated system. This poses some practical considerations. For example, force field param-  
48 eter optimisation may be necessary. For convergence of the predicted free energy change for a  
49 transformation, it is important that at each alchemical perturbation between two end states, the  
50 system is equilibrated to that intermediate state. Also, as perturbations need to be conservative,  
51 many windows may be needed to cope with the change between the two end states. As a result,  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 these simulations can become computationally intensive and access to parallel computing systems  
4 is required for thorough sampling regimes at realistic timescales.  
5  
6

7 In this study, MD simulations of  $\alpha\text{v}\beta\text{6}$  in complex with its natural ligand, TGF- $\beta\text{1}$ , enable  
8 us to investigate the dynamic and thermodynamic behaviour of ligand binding. We identify the  
9 contributions of active site residues to binding by monitoring how often they interact with the ligand.  
10 MD simulations are also performed for  $\alpha\text{v}\beta\text{6}$  in complex with each of the RGD mimetic derivatives.  
11 We investigate how different substituents and their enantiomers affect active site interactions of the  
12 potential antagonist with  $\alpha\text{v}\beta\text{6}$ . Furthermore, to calculate the difference in binding free energy  
13 between derivatives of the scaffold, FEP simulations have been performed on pairs of the RGD  
14 mimetics with each of the substituents.  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24

## 25 **Methods**

### 26 **Parameter Development**

27  
28  
29 All simulations were performed using the CHARMM force field.<sup>21</sup> The CHARMM General Force  
30 Field (CGenFF) is an extension which contains parameters for small drug-like molecules.<sup>22</sup> How-  
31 ever, not all atom types in the RGD mimetic, specifically the 1,2,3,4-tetrahydro 1,8-naphthyridine  
32 group, exist within CGenFF. Parameters consistent with the CHARMM force field have been de-  
33 veloped herein to enable the computational study of the potential  $\alpha\text{v}\beta\text{6}$  antagonists. Parameters  
34 were optimised according to the procedure outlined by Vanommeslaeghe et al.<sup>23,24</sup>  
35  
36  
37  
38  
39  
40  
41  
42  
43

44 The CHARMM force field consists of a collection of terms, which describe the potential energy  
45 of a system. Parameters in the force field, such as partial charges and equilibrium values of bond  
46 lengths, angles and dihedral angles, can be developed and optimised based on unique atom types.  
47 As not all atom types in our parent compound pre-existed within the force field, initial guesses  
48 for parameters were obtained using a CGenFF atom typing program.<sup>23,24</sup> Penalty scores were  
49 generated based on the level of optimisation required for these guess parameters. Basic validation  
50 is recommended for penalties between 10 and 50, and extensive optimisation is necessary for  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

penalty scores above 50. For computational efficiency, the scaffold was split into three fragments (Figure 3) and guess parameters were generated for each one. As fragment 1 was the only fragment to return penalty scores above 10, we concentrated our optimisation efforts on the partial charges and intramolecular parameters of the 1,8-tetrahydro naphthyridine group.

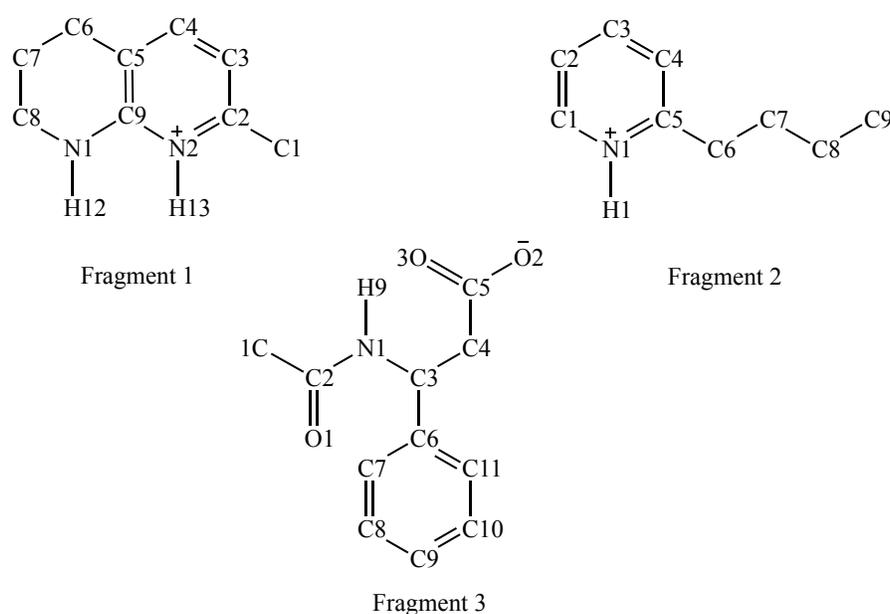


Figure 3: Fragments 1, 2 and 3, derived from the scaffold of the potential  $\alpha$ v antagonist. Atom labels correspond to those used in the parametrisation process.

Target data were generated by performing quantum mechanical (QM) calculations, using QChem.<sup>25</sup> Following Vanommeslaeghe et al., for each possible hydrogen bonding interaction, a complex was built of the MP2/6-31G\* optimised fragment and a single water molecule in the TIP3P geometry.<sup>26</sup> The water molecule was initially placed in an “ideal” position for hydrogen bonding and the interaction distance was optimised at the HF/6-31G\* level. QM values were compared with interaction distances and energies obtained from energy minimisations of the same complexes using the initial guess parameters. Partial charges were optimised so that model fragment-water interaction energies are within 0.2 kcal mol<sup>-1</sup> and distances are within 0.1 Å of the target data.

Bond length and angle parameters were optimised so that values from MP2/6-31G\* optimised fragments were replicated by energy minimisations using the force field. Penalty scores for initial

1  
2  
3 guess parameters indicated that, for fragment 1, one bond length and several bond angle parameters  
4 required attention. The penalty score for N1-C9 bond length parameters was 145. Therefore,  
5 the force constant and equilibrium terms were modified until the CHARMM energy minimised  
6 length was within 0.03 Å of the MP2/6-31G\* optimised structure. All bond angle parameters were  
7 optimised so that the angles were within 3° of the MP2/6-31G\* optimised structure.  
8  
9

10  
11  
12  
13 Due to the pseudo-planarity of fragment 1, the only parameters which required optimisation were  
14 those for the N1-C8-C7-C6 torsion. Optimisation of dihedral angle parameters used the MP2/6-  
15 31G\* potential energy surface (PES) as target data. The initial method employed to generate the  
16 QM surface involved an optimisation with the dihedral angle constrained and using the resulting  
17 conformation for the subsequent dihedral optimisation. However, this produced irregularities in the  
18 PES, due to the ring flip as the dihedral angle passed through 0°. To obtain a smooth PES, geometry  
19 optimisations were performed using a planar initial structure for each new dihedral constraint. The  
20 potential energy term describing rotation around a dihedral angle in the CHARMM force field  
21 features parameters for the force constant,  $K_\chi$ , the multiplicity,  $n$  and the phase  $\delta$ . The same  
22 scan was carried out using the CHARMM force field, with these dihedral parameters set to zero.  
23 A Monte-Carlo Simulated Annealing (MCSA) protocol<sup>27</sup> with exponential cooling was used to  
24 minimise the root mean square error between the QM and molecular mechanics (MM) energy  
25 profiles.  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38

39 To test how well existing CHARMM force field parameters describe the ligand derivative penalty  
40 scores for an aryl ring, with each of the substituents attached in turn, were obtained. From these  
41 scores, optimisation of partial charges on the CF<sub>3</sub> derivative was considered necessary. Dihedral  
42 angle parameters were also validated for the C-O-C-F torsion on the OCF<sub>3</sub> derivative. No further  
43 optimisation was needed for any other aryl substituents used in this study.  
44  
45  
46  
47  
48  
49  
50

## 51 Docking

52  
53  
54 Coordinates were taken from an X-ray crystal structure of a  $\alpha\text{v}\beta\text{6}$  dimer with the pro-domain of  
55 its natural ligand, TGF- $\beta\text{1}$ , bound (PDB code: 4UM9).<sup>10</sup> To prepare for docking, chains C, D  
56  
57  
58  
59  
60

1  
2  
3 and F were extracted from the equilibrated structure with all water molecules and ions removed  
4 (equilibration methodology outlined below). Using receptor generation software as part of the  
5 OpenEye docking toolkit,<sup>28,29</sup> Chain F (TGF- $\beta$ 1) was assigned as ligand and thus did not interact  
6 with the docked molecules. A box centred on ( $\beta$ 6)-Thr<sup>221</sup> with sides of length 21.0 x 22.7 x 27.3  
7 Å was situated to fully cover the TGF- $\beta$ 1 occupied binding site, giving a total receptor volume  
8 of 8680 Å<sup>3</sup>. Constraints were then applied, ensuring a metal chelate interaction with MIDAS and  
9 hydrogen bond donors to both of the carboxyl oxygen atoms on the ( $\alpha$ v)-Asp<sup>218</sup> residue.  
10  
11

12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Compounds were protonated according to physiological pH. The compounds studied were prepared using OMEGA,<sup>29</sup> for both (*R*) and (*S*) enantiomers. Conformers were generated using a truncated form of the MMFF94s force field,<sup>30</sup> a variant that excludes both Coulomb interactions and the attractive part of van der Waals interactions. A maximum energy difference of 20 kcal mol<sup>-1</sup> was allowed from the lowest energy conformer. These allowed molecules to explore additional conformational space. A maximum of 10,000 conformers per enantiomer was set and conformers within 0.5 Å of any others were considered duplicate and thus removed. Docking was performed using OpenEye FRED.<sup>28</sup> Compounds were docked using the high resolution setting with rotational and translational step sizes of 1 Å. Chemgauss4 was used to score the poses. The poses of each enantiomer were inspected for anomalies and the top scoring poses were chosen as the starting positions for MD simulations.

## Molecular Dynamics Simulations

Coordinates of chains C and D were used as starting structures for subunits  $\alpha$ v and  $\beta$ 6 respectively. Simulations involving the natural ligand used chain F coordinates. All water molecules and metal ions throughout the crystal structure were included. It is particularly important to retain crystallographic water molecules for FEP simulations as they stabilise the system and improve the equilibration process.<sup>20</sup> In accordance with physiological pH, the zwitterion form of the RGD mimetic was used, with the 1,2,3,4-tetrahydro 1,8-naphthyridine protonated and the carboxyl group deprotonated. The protonated states of arginine and lysine residues were used and all aspartic and

1  
2  
3 glutamic acids were deprotonated. Histidine residues were treated as neutral, with the nitrogen atom  
4 nearer the backbone protonated. Procedures to build hydrogen atoms, solvate the system and apply  
5 periodic boundary conditions were generated using the quick MD simulator module in CHARMM-  
6 GUI.<sup>31</sup> The system was solvated in a truncated octahedral periodic boundary cell with edge distances  
7 of 10 Å to construct an explicitly modelled solvent consisting of 16,886 TIP3P water molecules,<sup>26</sup>  
8 eight Mg<sup>2+</sup> and four Cl<sup>-</sup> ions, to give a net neutral charge. The concentration of the counter ions  
9 matched conditions used in cell adhesion assays performed on these complexes. To optimise the  
10 solvent positions, all heavy atoms were fixed, except for water molecules, during 50 steps of steepest  
11 descent (SD) and 50 steps of Adopted Basis Newton-Raphson (ABNR) minimisation. Protein and  
12 metal ion parameters were obtained from the C36 version of the CHARMM force field.<sup>32</sup> Parameters  
13 for metal ions were developed and validated by Beglov et al.<sup>33</sup> and are commonly employed in  
14 biomolecular studies.

15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27 Upon system setup, the NAMD software<sup>34</sup> was used for simulations of all complexes. Firstly,  
28 the solvated crystal structure, still containing TGF-β1, was minimised for 20 ps using a conjugate  
29 gradient and line search algorithm. Protein backbone and side chain restraints were applied using  
30 harmonic constraints with force constants of 10 kcal mol<sup>-1</sup> Å<sup>-2</sup> and 5 kcal mol<sup>-1</sup> Å<sup>-2</sup>. The system  
31 was heated to 298 K in increments of 3 K every 1 ps using NAMD velocity reassignment. Backbone  
32 and side chain restraints were gradually switched off during a 2 ns equilibration period in the NVT  
33 ensemble. The coordinates of the equilibrated receptor were used for ligand docking.

34  
35  
36  
37  
38  
39  
40  
41 To ensure canonical interactions were maintained during an additional 1 ns equilibration of the  
42 docked complexes, both carboxyl oxygen atoms on the aspartate mimetic were constrained to a  
43 distance of 2 Å from the Mg<sup>2+</sup> ion. The polar hydrogen atoms on the protonated 1,2,3,4-tetrahydro  
44 1,8-naphthyridine segment were constrained to a distance of 2 Å from the carboxyl oxygen atoms on  
45 (αv)-Asp<sup>218</sup>. An initial force constant of 10 kcal mol<sup>-1</sup> Å<sup>-2</sup> was used for all distance constraints.  
46 Force constants for distance constraints were steadily decreased to 2.5 kcal mol<sup>-1</sup> Å<sup>-2</sup> during  
47 an equilibration of 0.5 ns in the NVT ensemble and 0.5 ns in the NPT ensemble. This meant  
48 all canonical interactions were present at the start of production runs. Five independent 10 ns  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 production runs were performed in the NPT ensemble for the (*R*)- and (*S*)-enantiomers of each  
4 derivative. Simulations of  $\alpha v\beta 6$  bound with TGF- $\beta 1$  were also performed, resulting in a total of  
5 65 simulations of 10 ns. Temperature was controlled using Langevin dynamics parameters, with  
6 a friction coefficient of  $5 \text{ ps}^{-1}$  for all equilibration and production runs. Constant pressure was  
7 maintained using the Langevin piston Nosé-Hoover method<sup>35</sup> with a target pressure of 1 atm. A  
8 cutoff distance of  $12 \text{ \AA}$  was used for van der Waals pairs, with a switching function at a distance of  
9  $10 \text{ \AA}$ . The electrostatic potential energy was computed using the Particle Mesh Ewald method.<sup>36</sup>  
10 The SHAKE algorithm<sup>37</sup> was used to fix all bond lengths involving hydrogen atoms and a timestep  
11 of 2 fs was used. Upon completion of production runs, all solvent molecules were removed except  
12 those within  $10 \text{ \AA}$  of the ligand, in order to expedite the analysis. Trajectories were sampled every  
13 20 ps, resulting in 500 frames for each replica.  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

## 26 27 **Free Energy Perturbation Simulations**

28  
29 FEP simulations involve an alchemical transformation between two structurally related ligands.<sup>19</sup>  
30 A change is made to the system so that the potential energy is equivalent to the original potential  
31 energy with an additional “perturbing” potential energy term ( $V_{BA}$ ):  
32  
33  
34  
35  
36

$$37 \quad U_B = U_A + V_{BA} \quad (2)$$

38  
39 where  $U_A$  and  $U_B$  are the respective potential energies of state A and B, which represent the states  
40 where a different ligand is bound in each. States A and B should arise from the same conformational  
41 space and therefore,  $V_{BA}$  needs to be very small. To address this, the transformation of A to B is  
42 divided into several discrete simulations, called windows, which are connected using a coupling  
43 parameter,  $\lambda$ .  
44  
45  
46  
47  
48  
49

50  
51 Figure 4 shows the thermodynamic cycle of a transformation of the hydrogen derivative of the  
52 RGD mimetic (state A) into the methoxy derivative (state B).  $\Delta G_1$  and  $\Delta G_2$  are the binding free  
53 energies of each ligand, while  $\Delta G_3$  and  $\Delta G_4$  are the free energies of transforming one ligand into  
54  
55  
56  
57  
58  
59  
60

the other, as the free ligand and in complex with the receptor. At each end point of the simulation (ligand A and ligand B)  $\lambda$  equals zero or one. At intervals between these end points, bonded and non-bonded parameters are scaled so that they are “switched off” for outgoing atoms and “switched on” for incoming atoms. There are two ways to alchemically mutate ligands. The single topology approach<sup>38</sup> uses dummy atoms for appearing/disappearing atoms, while gradually scaling parameters so that relevant atoms of the initial ligand are turned off and they become dummy atoms. The nature of this method means that, although transformations converge better, it cannot cope with large changes in ligands. In the dual topology approach,<sup>39</sup> the topology of the mutating species contains both ligands simultaneously. Intermolecular interactions are turned on for one ligand as they are turned off for the other. Typically, a dual topology method is easier to use and can cope with larger changes in ligands compared to the single topology method.<sup>40</sup> In our FEP simulations, we employ the dual topology approach.

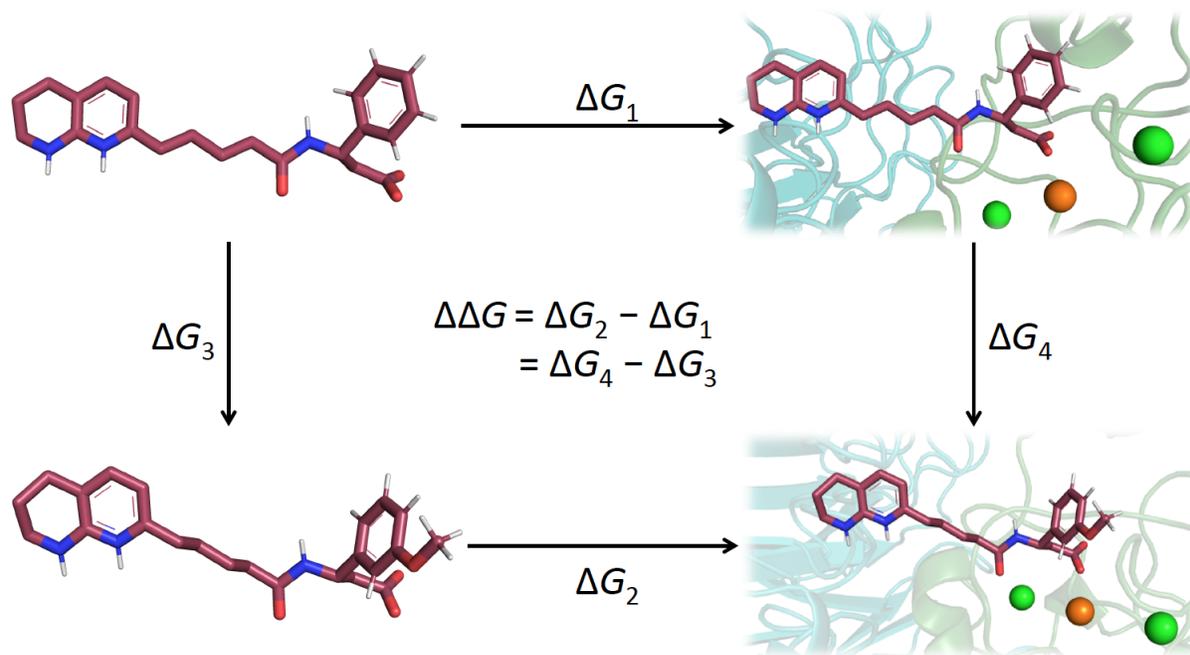


Figure 4: Thermodynamic cycle describing the binding of two ligands.  $\Delta G_1$  and  $\Delta G_2$  represent the free energy of binding of two ligands.  $\Delta G_3$  and  $\Delta G_4$  describe the free energy change of an alchemical transformation of one ligand into the other, as the free ligand in solution and bound to the receptor, respectively. The relative free energy of binding of two ligands is the difference between the alchemical free energy changes.

1  
2  
3 A soft-core potential is used to avoid “end-point catastrophes”, where disappearing atoms can  
4 leave empty pockets and appearing atoms clash with the existing environment.<sup>41,42</sup> This eliminates  
5 the singularities at each end point by progressively scaling interactions of outgoing atoms and  
6 incoming atoms. The short range repulsive term in the standard Lennard-Jones potential is scaled to  
7 allow “soft” overlap of van der Waals spheres at regions surrounding incoming and outgoing atoms.  
8 The Coulombic term in the potential is also scaled to avoid abnormal electrostatic interactions  
9 between the softened atoms and their environment. Electrostatic interactions of outgoing atoms  
10 are decoupled from the system over the  $\lambda$  range of 0 to 0.5, while the electrostatics for incoming  
11 atoms are coupled to the system over the  $\lambda$  range of 0.5 to 1.  
12  
13  
14  
15  
16  
17  
18  
19  
20

21 To calculate relative binding free energies of pairs of ligands, ten different alchemical transfor-  
22 mations of aryl substituents were performed. For each perturbation, three replicas of the forwards  
23 and backwards transformation were performed, resulting in 60 FEP simulations. Ahead of system  
24 setup, dual topologies were constructed for both the free and bound ligands. To prevent system  
25 drift, caused by the charged ends of the free ligand coming together, the transformations of the free  
26 ligand in solvent started from structures close to an energy minimised conformation. The setup  
27 of systems and free energy simulations were performed using the procedures for the standard MD  
28 simulations described above. During FEP,  $\lambda$  was increased from 0.0 to 0.1 in 16 discrete steps,  
29 from 0.1 to 0.9 in steps of 0.02 and from 0.9 to 1.0 in 16 steps to give a total of 72 windows for  
30 each transformation. Equilibration was performed for 20 ps at the start of each window, followed  
31 by 100 ps of sampling.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44

## 45 **Results and Discussion**

### 46 **Parameter Development**

47  
48  
49 For the 1,2,3,4-tetrahydro 1,8-naphthyridine group (fragment 1 in Figure 3), Table 1 shows the  
50 optimised interaction energies and distances for hydrogen bonding interactions with polar atoms  
51 on the antagonist at the HF/6-31G\* level, and when using the CHARMM force field once partial  
52  
53  
54  
55  
56  
57  
58  
59  
60

charges were optimised. Interaction energies are within  $0.2 \text{ kcal mol}^{-1}$ , as recommended by parameter optimisation methodology.<sup>22</sup> Although the interaction distance between the fluorine atoms on the  $\text{CF}_3$  derivative and the TIP3P oxygen atom is above the recommended  $0.1 \text{ \AA}$ , similar disagreement has been acceptable in other major parameterisation efforts.<sup>22</sup>

Table 1: QM and MM interaction energies and distances of a single water molecule in a hydrogen bonding interaction with polar atoms on fragment 1 and the  $\text{CF}_3$  aryl derivative. O represents the oxygen atom on the TIP3P water molecule. Distances are in  $\text{\AA}$  and energies are given in  $\text{kcal mol}^{-1}$ .

Interaction	HF/6-31G*		Optimised CHARMM Parameters		Difference	
	Distance	Energy	Distance	Energy	Distance	Energy
H12 - O	2.1	-16.4	2.0	-16.4	0.1	0.0
H13 - O	2.0	-16.2	1.9	-16.4	0.1	0.2
CF - O	2.3	-1.9	2	-1.9	0.3	0.0

Figure 5 shows the MP2/6-31G\* bond lengths and partial charges assigned to the methylated 1,8-naphthyridine fragment. The similarity in bond lengths between the central carbon atom and adjacent nitrogen atoms suggests that the charge delocalisation extends across the ring to include both nitrogen atoms, rather than a localised aromatic pyridine-piperidine structure. This is reflected in the partial charges assigned to these atom types in the developed force field parameters (full parameters can be found in the Supporting Information).

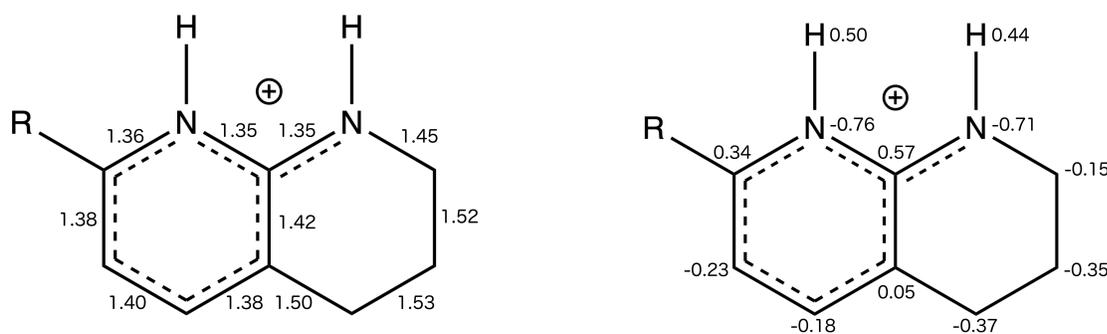


Figure 5: Bond lengths and atomic partial charges calculated at the MP2/6-31G\* level for fragment 1. Bond lengths are shown in  $\text{\AA}$ .

Optimised values for the high penalty bond lengths and angles in fragment one are shown in Table 2. These values are also in close agreement with the bond lengths and angles in four naphthyridine

containing molecules<sup>43-46</sup> from the Cambridge Crystallographic Data Centre (CCDC)<sup>47</sup> (Table S1 in Supporting Information).

Table 2: Bond lengths and angles on fragment 1 from MP2/6-31G\* geometry optimisation and energy minimisation using optimised CHARMM force field parameters. Bond lengths shown in Å and bond angles shown in degrees.

Bond	MP2/6-31G* Optimisation	Optimised CHARMM Parameters	Difference
N1 - C9	1.34	1.33	0.01
N1 - C9 - N2	119	118	1
C9 - N1 - H12	120	120	0
C9 - N1 - C8	122	119	3
C5 - C9 - N1	123	124	1

Figure 6 shows the potential energy surface of the N1-C8-C7-C6 dihedral angle, calculated using different methods. There is a closer match between MM and QM surfaces once dihedral parameters,  $K$ ,  $n$  and  $\delta$ , are optimised using a MCSA protocol. The penalty score returned for the C-O-C-F dihedral parameters in the OCF<sub>3</sub> derivative was 98, implying necessity for optimisation. However, the QM and MM PES around the torsion indicated a sufficient match between the curves when using the initial guess parameters; therefore, no optimisation was performed.

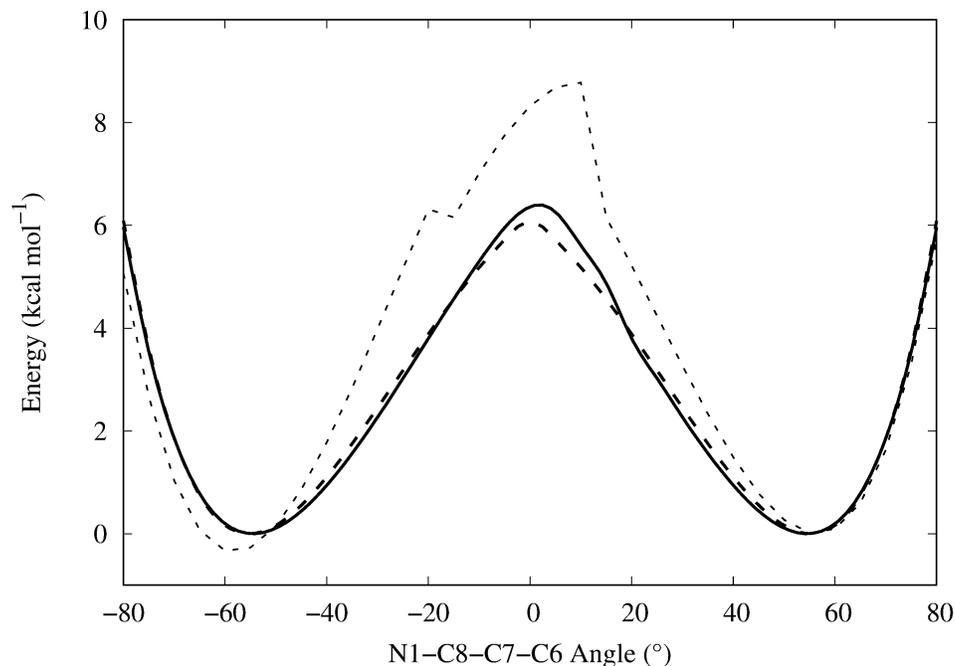


Figure 6: Dihedral scan around the N1-C8-C7-C6 angle showing the MP2/6-31G\* surface (dashed line), the CHARMM initial guess (dotted line) and the CHARMM surface after parameters are optimised (solid line).

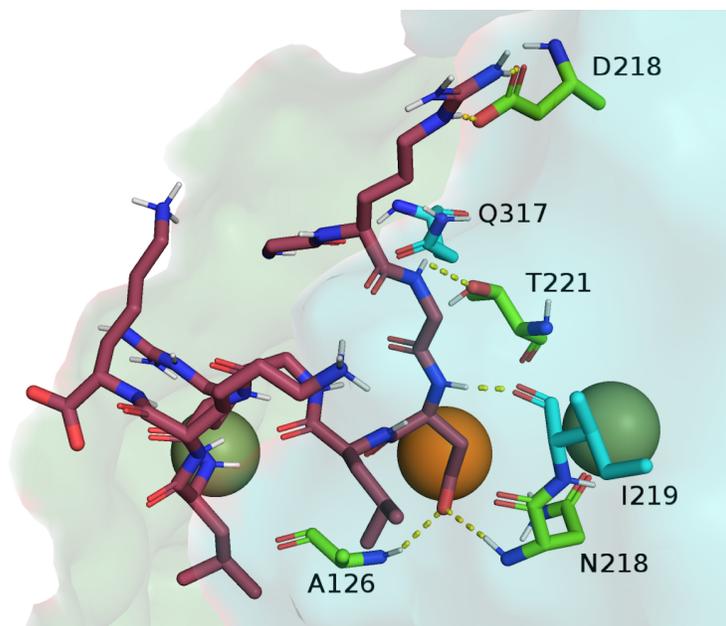
## Molecular Dynamics Simulations

### Protein-Ligand Interactions of TGF- $\beta$ 1

Polar residues that are important for binding can be identified by investigating the dynamics of hydrogen bonds between the natural ligand of  $\alpha$ v $\beta$ 6, TGF- $\beta$ 1, and the receptor. From the crystal structure, it is clear that ( $\alpha$ v)-Asp<sup>218</sup>, ( $\beta$ 6)-Ala<sup>126</sup>, ( $\beta$ 6)-Asn<sup>218</sup> and ( $\beta$ 6)-Thr<sup>221</sup> form hydrogen bonds with TGF- $\beta$ 1. The fraction of time that these interactions were maintained over the 50 ns MD simulation was measured. A pair of atoms are considered to be hydrogen bonded if a polar hydrogen atom is within 2.5 Å of an oxygen, nitrogen or fluorine atom. A bidentate interaction between ( $\alpha$ v)Asp<sup>218</sup> and the Arg residue of the RGD unit on the ligand was present for 82% of the simulation time. ( $\beta$ 6)-Ala<sup>126</sup>, ( $\beta$ 6)-Asn<sup>218</sup> and ( $\beta$ 6)-Thr<sup>221</sup> were all hydrogen bonded with TGF- $\beta$ 1 for over 90% of the simulation time. The MD simulation indicated two hydrogen bonds formed between the Asp residue of the RGD binding tripeptide in the ligand and ( $\beta$ 6)-Asn<sup>218</sup>. One

1  
2  
3 interaction involved the amide backbone of the ( $\beta 6$ )-Asn<sup>218</sup> residue for 98% of the simulation time  
4 and the other involved the side chain of ( $\beta 6$ )-Asn<sup>218</sup> for 66% of the simulation time. The hydrogen  
5 bond with the side chain is not present in the starting structure and is only observed as the result  
6 of dynamics. Both interactions with ( $\beta 6$ )-Asn<sup>218</sup> are known among RGD ligands and have been  
7 previously recognised.<sup>10,48,49</sup> The metal chelate interaction between Asp<sup>RGD</sup> and the Mg<sup>2+</sup> MIDAS  
8 ion was maintained throughout the entire simulation.  
9  
10  
11  
12  
13  
14

15 Figure 7 depicts the position of the binding site residues that form hydrogen bonds with TGF-  
16  $\beta 1$ , identified by the crystal structure and MD simulations. Residues ( $\beta 6$ )-Ile<sup>219</sup> and ( $\beta 6$ )-Gln<sup>317</sup>  
17 are also shown. These do not appear to interact in the crystal structure but are in close proximity  
18 with TGF- $\beta 1$  for 78% and 7% of the MD simulation time, respectively. This highlights how  
19 MD simulations can identify active site residues that contribute to ligand binding, which are not  
20 observed in a static crystal structure. In an ongoing study, these receptor conformations generated  
21 by MD simulations are being used for docking as a way to extend the receptor conformational  
22 search space and identify further antagonists.  
23  
24  
25  
26  
27  
28  
29  
30  
31



32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52 Figure 7: Residues that form interactions with TGF- $\beta 1$  (maroon backbone) over a 50 ns MD  
53 simulation. Interacting residues in the crystal structure are shown (light green) as well as residues  
54 that do not interact in the crystal structure, but gain an interaction during MD (light blue). The  
55 Mg<sup>2+</sup> MIDAS ion (orange) and Ca<sup>2+</sup> ions (dark green) are also shown.  
56  
57  
58  
59  
60

1  
2  
3 Hydrophobic interactions are also important for the binding of TGF- $\beta$ 1 to  $\alpha$ v $\beta$ 6. The 244-  
4 LGRLK-248 sequence directly following the RGD motif folds into an amphipathic  $\alpha$ -helix, which  
5 fits into a  $\beta$ 6-specific hydrophobic pocket, as previously reported.<sup>10,12,50,51</sup> Leu<sup>224</sup> and Leu<sup>247</sup> form  
6 lipophilic contacts with this pocket, which is formed by residues distinct to  $\beta$ 6 when compared to  
7 other RGD binding members of the integrin family such as  $\alpha$ v $\beta$ 3 and  $\alpha$ v $\beta$ 5.<sup>10</sup> Contact between the  
8 amphipathic  $\alpha$ -helix of TGF- $\beta$ 1 and the hydrophobic pocket of  $\beta$ 6 was maintained throughout the  
9 MD simulations.  
10  
11  
12  
13  
14  
15  
16  
17

### 18 **Protein-Ligand Interactions of RGD Mimetics**

19  
20  
21 By monitoring the dynamics of hydrogen bonding interactions between each derivative of the RGD  
22 mimetic and the receptor, we characterise the nature of ligand binding. The stability of the canonical  
23 interactions could reflect the potency of each ligand, as an active compound should maintain these  
24 interactions and remain bound in the active site. In principle, there could be a link between  
25 analogues with higher pIC<sub>50</sub> values and better maintained canonical interactions. All derivatives  
26 of the pseudo-RGD compound remain bound throughout the MD simulations, through at least  
27 one component of the canonical interactions. Table 3 shows the pIC<sub>50</sub> values of each analogue,  
28 which are a measure of activity<sup>5</sup> (with additional pharmacological data available in reference 5)  
29 and the fraction of time that the canonical interactions were maintained. The interaction frequency  
30 is the proportion of frames with the interaction present with respect to the total number of frames,  
31 averaged over five 10 ns simulations. The bidentate interaction is present in all derivatives of the  
32 RGD mimetic to a varying extent. However, there is no correspondence between the stability of  
33 the bidentate interaction and pIC<sub>50</sub>.  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Table 3: Interaction frequency in % of simulation time that the bidentate,  $\text{Mg}^{2+} - \text{O}^1$  and  $\text{Mg}^{2+} - \text{O}^2$  interactions are maintained for each analogue of the RGD mimetic. Contact frequencies between atoms on the aryl rings and active site residues are also shown in columns 8-10. The experimental activity of each analogue is shown by its  $\text{pIC}_{50}$  value.<sup>5</sup> The molecular docking score (arbitrary units), generated using OpenEye FRED<sup>28</sup> and OpenEye OMEGA,<sup>29</sup> is also shown.

Substituent	Enantiomer	$\text{pIC}_{50}$	Docking Score	Interaction Frequency (% Time)					
				Bidentate Interaction	$\text{Mg}^{2+} - \text{O}^1$	$\text{Mg}^{2+} - \text{O}^2$	( $\beta 6$ )-Ser <sup>127</sup>	( $\beta 6$ )-Ala <sup>126</sup>	( $\beta 6$ )-Ile <sup>219</sup>
H	S	5.7	-9.99	89	100	0	96	8	20
	R		-9.19	96	100	100	24	24	0
F	S	6.1	-9.64	96	100	46	50	22	2
	R		-9.29	29	89	80	39	23	5
CH <sub>3</sub>	S	6.4	-7.96	100	100	73	81	38	49
	R		-4.64	100	100	100	47	43	2
OCH <sub>3</sub>	S	6.5	-9.38	65	100	100	100	62	0
	R		-6.57	100	100	100	84	3	0
OCF <sub>3</sub>	S	6.7	-8.49	98	100	40	45	19	0
	R		-3.59	99	100	0	0	0	3
CF <sub>3</sub>	S	7.1	-9.33	74	100	100	57	43	25
	R	5.2	-5.70	19	100	100	0	0	3

The loss of the bidentate interaction is caused by a slight overall translation of the ligand in the binding site so that only one polar hydrogen atom on the naphthyridine fragment is within 2.5 Å of ( $\alpha$ v)-Asp<sup>218</sup>. Although the bidentate interaction is lost for a considerable amount of time in some ligands, a monodentate interaction between the Arg mimetic and ( $\alpha$ v)-Asp<sup>218</sup> is commonly observed for all ligands (Table S2 in the Supplementary Information). All analogues of the ligand form at least one metal chelate interaction with the  $\text{Mg}^{2+}$  ion throughout the MD simulations. The difference between interactions  $\text{Mg}^{2+} - \text{O}^1$  and  $\text{Mg}^{2+} - \text{O}^2$  is the oxygen atom on the carboxyl group that the  $\text{Mg}^{2+}$  interacts with. Given the proximity of the chiral centre to the MIDAS site, it might be expected that interaction frequencies for the less active enantiomer with MIDAS amino-acid residues will be lower as the calculations suggest. Given the conformational flexibility of the molecule however, what is less expected is the impact on the bidentate interaction frequency between ( $\alpha$ v)-Asp<sup>218</sup> and the 1,8-naphthyridine at the other end of the molecule which is much lower for the less active enantiomer compared to the more active enantiomer (R-CF<sub>3</sub> 19% and S-CF<sub>3</sub> 74%). Nevertheless, the interaction with the MIDAS ion is better maintained than the bidentate hydrogen bond in all cases. Therefore, we observe that the stability of the metal chelate interaction between Asp<sup>RGD</sup> and the  $\text{Mg}^{2+}$  ion is more important than the bidentate hydrogen bonding to ( $\alpha$ v)-Asp<sup>218</sup> for ligand binding.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

The position of the aryl ring on the bound ligand is indicated in Figure 8. The metal chelate interaction formed by the carbonyl and  $\text{Mg}^{2+}$  MIDAS ion points inwards into the binding pocket, forcing the aryl ring to become solvent exposed. Nearby residues are on the  $\beta 1$ - $\alpha 1$  and  $\alpha 2$ - $\alpha 3$  loops, specifically  $(\beta 6)$ -Ala<sup>126</sup>,  $(\beta 6)$ -Ser<sup>127</sup>,  $(\beta 6)$ -Asn<sup>218</sup>,  $(\beta 6)$ -Ile<sup>219</sup>,  $(\beta 6)$ -Asp<sup>220</sup> and  $(\beta 6)$ -Thr<sup>221</sup>. The side chain of  $(\beta 6)$ -Ala<sup>126</sup> and the backbone of  $(\beta 6)$ -Asn<sup>218</sup> contribute to the  $\beta 6$  hydrophobic binding pocket.<sup>10</sup> All (*S*)-enantiomer derivatives of the ligand show an interaction with both groups (Table S2). Further contacts within 2.5 Å of each substituted ring and any atom on the  $\beta 6$  unit were investigated. Table 3 shows the frequency of contacts with  $(\beta 6)$ -Ala<sup>126</sup>,  $(\beta 6)$ -Ser<sup>127</sup> and  $(\beta 6)$ -Ile<sup>219</sup>. From the proximity of the ring to  $(\beta 6)$ -Ser<sup>127</sup>, we observe that the (*S*)-H, (*S*)-CH<sub>3</sub> and (*S*)-OCH<sub>3</sub> derivatives remain in close contact with the  $\beta 1$ - $\alpha 1$  loop. The remaining *S*-derivatives, (*S*)-CF<sub>3</sub>, (*S*)-F and (*S*)-OCF<sub>3</sub>, stay close to the  $\beta 1$ - $\alpha 1$  loop, but for less than 50% of the simulation time. As no other contacts are formed between these derivatives and the receptor for a significant amount of time, we suggest that solvent interactions with these more polar substituents influence the position of the aryl ring.

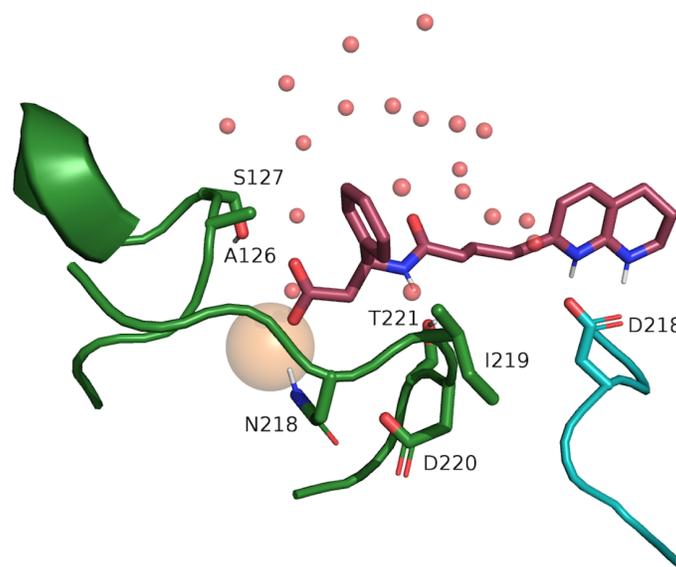


Figure 8: Position of the aryl ring on the (*S*)-RGD mimetic within the binding site. Nearby residues on the  $\beta 1$ - $\alpha 1$  and  $\alpha 2$ - $\alpha 3$  loops (green) are labelled. The oxygen atoms of water molecules within 7 Å of the ring are also shown (red spheres).

1  
2  
3 MD simulations of the natural ligand, TGF- $\beta$ 1, indicated that residues ( $\beta$ 6)-Ala<sup>126</sup>, ( $\beta$ 6)-  
4 Thr<sup>221</sup> and ( $\beta$ 6)-Asn<sup>218</sup> in the receptor are all involved in binding. Therefore, hydrogen bonding  
5 interactions between these residues and the RDG mimetics were investigated. The carboxyl oxygen  
6 atoms on the (*S*)-forms of the H, F, CH<sub>3</sub> and OCF<sub>3</sub> derivatives interacted with ( $\beta$ 6)-Ala<sup>126</sup> for  
7 67%, 45%, 10% and 60% of the simulation time, respectively. No stable interactions are identified  
8 between any derivatives of the ligand and ( $\beta$ 6)-Thr<sup>221</sup>. Although interactions between the ligands  
9 and residue ( $\beta$ 6)-Thr<sup>221</sup> are present in the docked structures, they are lost during equilibration of the  
10 systems and are not observed during dynamics. In the docked structures, ( $\beta$ 6)-Thr<sup>221</sup> interacts with  
11 the carbonyl oxygen atom on the amide backbone of the ligand. However, during equilibration, the  
12 carboxyl group rotates away from the receptor to become more solvent exposed.  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24

### 25 **Cation- $\pi$ interactions in $\alpha$ v $\beta$ 6-RGD systems**

26  
27 A cation- $\pi$  interaction consists of an electrostatic attraction between an electron rich aryl ring and  
28 an electron deficient cation. This type of interaction is common in proteins; an early data mining  
29 study of a subset of the PDB found that over 70% of all Arg side chains are near an aromatic side  
30 chain.<sup>52</sup> The crystal structure of  $\alpha$ v $\beta$ 6 - TGF- $\beta$ 1 is no exception, with a face-to-face intermolecular  
31 cation- $\pi$  interaction between Arg<sup>RGD</sup> and residue ( $\alpha$ v)-Tyr<sup>178</sup> (Figure 9). A mixture of factors  
32 govern cation- $\pi$  interaction orientation, such as competitive hydrogen bonding and the influence of  
33 solvent.<sup>53</sup> However, this parallel geometry is more common in proteins than T-shaped geometries,  
34 although T-shaped geometries are preferred in the gas-phase.<sup>52,54</sup> To monitor the persistence of  
35 the cation- $\pi$  interaction during MD, the distance between the cationic Arg group (the guanidinium  
36 carbon in Arg) in TGF- $\beta$ 1 and the centre of the aromatic ( $\alpha$ v)-Tyr<sup>178</sup> ring was measured. A distance  
37 less than 5 Å, a common geometrical constraint used in previous studies, indicated an interaction.  
38 The average distance between the two residues was 4.41 Å with a standard deviation of 0.48 Å. The  
39 face-to-face geometry was also maintained throughout.  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



Figure 9: Cation- $\pi$  interactions between the cationic nitrogen atoms on Arg<sup>RGD</sup> (left) or the 1,2,3,4-tetrahydro 1,8-naphthyridine group of the (*S*)-CF<sub>3</sub> derivative (right) and the aromatic side chain of ( $\alpha$ v)-Tyr<sup>178</sup>.

As the overarching aim is to design a compound that competitively binds with  $\alpha$ v $\beta$ 6, it would be sensible to take advantage of the position of ( $\alpha$ v)-Tyr<sup>178</sup> and replicate this cation- $\pi$  interaction with the RGD mimetics. The cation in this case is the protonated 1,2,3,4-tetrahydro 1,8-naphthyridine group. Although the interaction between the naphthyridine group, an aromatic cation, and the aromatic side chain of ( $\alpha$ v)-Tyr<sup>178</sup> could be dominated by  $\pi$ - $\pi$  interactions, a study by Tsuzuki et al.<sup>55</sup> found, through high level ab initio calculations, that interactions of benzene complexes with aromatic cations should be categorised as cation- $\pi$  interactions as they are stabilised by large electrostatic and induction interactions. The distance between the positively charged nitrogen atom and the centre of the aromatic ( $\alpha$ v)-Tyr<sup>178</sup> ring was measured for each of the RGD derivatives. Table 4 shows the distances, averaged over 10 ns of MD. The majority of compounds had an average distance within the 5 Å cutoff. The cation- $\pi$  interaction orientation sampled most was a face-to-face geometry of the ( $\alpha$ v)-Tyr<sup>178</sup> side chain with the naphthyridine group (Figure 9). Thus, the RGD mimetics replicate this type of interaction seen with TGF- $\beta$ 1, in addition to the canonical hydrogen bonding interactions.

Table 4: The distance between the protonated N atom on each RGD derivative and the centre of the aromatic ring on ( $\alpha$ v)-Tyr<sup>178</sup>. Distances have been averaged over 10 ns of MD simulations, with the standard deviation shown.

Substituent	NH <sup>+</sup> - Tyr <sup>178</sup> Distance (Å)
( <i>S</i> )-H	3.77 ± 0.33
( <i>R</i> )-H	5.01 ± 0.86
( <i>S</i> )-F	4.80 ± 0.41
( <i>R</i> )-F	4.78 ± 1.06
( <i>S</i> )-CH <sub>3</sub>	4.63 ± 0.43
( <i>R</i> )-CH <sub>3</sub>	4.13 ± 0.50
( <i>S</i> )-OCH <sub>3</sub>	3.82 ± 0.47
( <i>R</i> )-OCH <sub>3</sub>	4.21 ± 0.49
( <i>S</i> )-OCF <sub>3</sub>	5.21 ± 0.44
( <i>R</i> )-OCF <sub>3</sub>	5.33 ± 0.69
( <i>S</i> )-CF <sub>3</sub>	4.19 ± 0.62
( <i>R</i> )-CF <sub>3</sub>	4.33 ± 0.50

## Free Energy Perturbation Simulations

To assess the convergence of the alchemical calculations, a method outlined by Klimovich et al<sup>56</sup> was used. In our example (Figure 10) we consider the transformation of a hydrogen substituent to an OCF<sub>3</sub> substituent, with the ligand in the bound state. This transformation was chosen as it is the most diverse in terms of the number and types of atoms perturbed and therefore, it might be most likely to have convergence issues. The free energy change is calculated using an increasing fraction of the simulation data (i.e. 0-0.1, 0.0-0.2). Also plotted in Figure 10 is the free energy change calculated with the reverse proportion of the data (i.e. 0.9-1.0, 0.8-1.0). Both sets of data remain within error of the final value and so the calculation is considered converged. Furthermore, the relative free energies of the forwards and backwards transformations were compared, as the values are expected to be identical but of opposite sign. The H → OCF<sub>3</sub> perturbation of the free ligand gave a relative free energy change of -8.9 kcal mol<sup>-1</sup>, while the OCF<sub>3</sub> → H perturbation resulted in a relative free energy change of 8.5 kcal mol<sup>-1</sup>. The same perturbations gave relative free energy changes of -10.4 kcal mol<sup>-1</sup> and 9.8 kcal mol<sup>-1</sup> when the ligand was bound. The relative free energies reflect that, due to its stronger non-covalent interactions, the OCF<sub>3</sub> derivative is favoured

over the hydrogen substituent in both solvent and in complex with the receptor. The similarity for the forwards and backwards transformation in both states suggests that conditions are met for the convergence of the alchemical simulations.

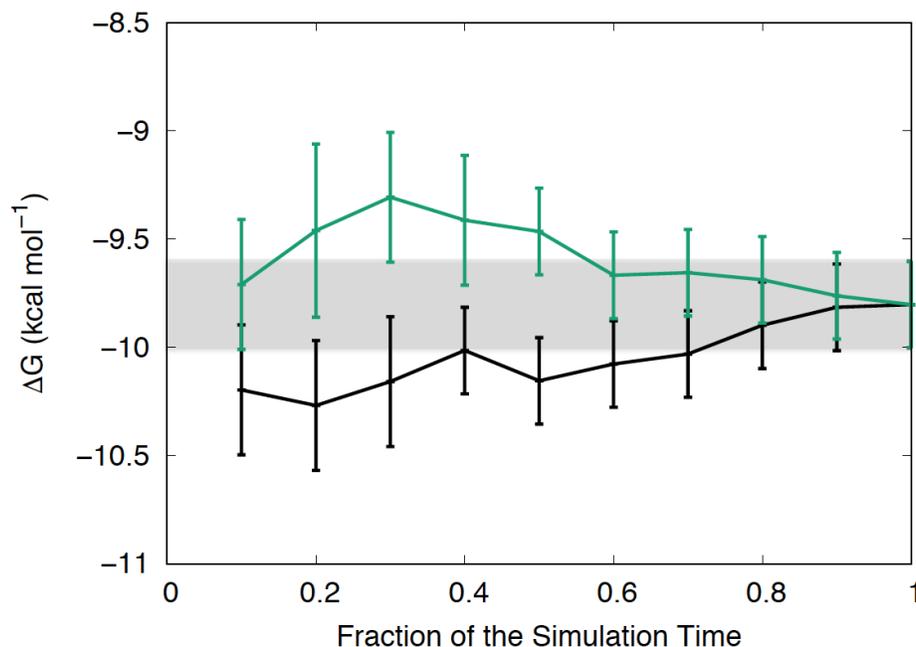


Figure 10: Convergence assessment of the transformation of the hydrogen derivative to the  $\text{OCF}_3$  derivative, with the ligand in the bound state. The forward (black line) and the reverse (green line) simulation time series are shown. The horizontal grey strip indicates the equilibrated region.

Calculated relative free energies of binding are shown in Table 5. The  $\Delta G$  values for the transformations of the free ligand and in complex with the receptor were evaluated with the Bennett Acceptance Ratio (BAR) method,<sup>57</sup> using the ParseFEP tool in VMD.<sup>58</sup>  $\Delta\Delta G$  values were obtained by taking the difference between each of these transformations. These theoretical values are compared with  $\Delta\Delta G$  values obtained from the experimental  $\text{pIC}_{50}$  values. For all ten alchemical transformations, the sign of the relative free energy difference is correctly predicted, although the magnitude is almost always over-estimated but in a non-systematic manner. As all derivatives, with the exception of  $\text{CF}_3$ , were measured as racemic mixtures, there is a probable  $\pm 0.3$  log error on each experimental binding free energy, which should also be taken into account. The calculated values are in within  $1.5 \text{ kcal mol}^{-1}$  of experiment, with the exception of the  $\text{H} \rightarrow$

1  
2  
3 CH<sub>3</sub>, CH<sub>3</sub> → OCH<sub>3</sub> and OCH<sub>3</sub> → OCF<sub>3</sub> aryl transformations. In order to compare the results of  
4 the FEP simulations with simpler methods, we investigate the relationship between docking scores  
5 and experimental activity (Table 3). Upon docking of the ligands into the equilibrated receptor,  
6 a score is obtained which reflects the quality of the docking of the RGD mimetic to the αvβ6  
7 binding site. Although the values are somewhat empirical and cannot be directly related to binding  
8 affinity, docking scores can be useful for ranking compounds or distinguishing between active and  
9 non-active compounds.<sup>59</sup> The docking scores shown in Table 3 do not correlate with the pIC<sub>50</sub>  
10 values. Despite this, the scores do consistently indicate that the (*S*)-enantiomers will be more active  
11 than the (*R*)-enantiomers, as expected. As the relative free energies obtained for the ligands agree  
12 more closely with experiment, compared to the docking scores, we suggest that a more physically  
13 realistic model has resulted from FEP simulations for this system, albeit at higher computational  
14 cost.

15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27 The narrow range of activity exhibited by these ligands means it is difficult to rank the derivatives,  
28 but is not atypical of many such examples in the field.<sup>60</sup> The close match between the values for each  
29 of the derivatives shows that FEP calculations model the system well, giving a solid foundation for  
30 the associated analysis. While the equilibrium MD simulations provide information on the dynamic  
31 behaviour of interacting residues (as detailed in Table 3), FEP simulations provide a rigorous  
32 way to quantify various physical factors, such as changes in hydrophobic and hydrogen bonding  
33 interactions. Taking the transformation from the hydrogen derivative to the OCF<sub>3</sub> derivative as an  
34 example, the more lipophilic OCF<sub>3</sub> substituent is expected to have more favorable interactions with  
35 the hydrophobic binding pocket in the β6 subunit. This is reflected by the experimental free energy  
36 of binding, which shows a difference of -1.4 kcal mol<sup>-1</sup> between the two derivatives. Calculated  
37 values show a similar relative binding free energy of -1.9 kcal mol<sup>-1</sup>. The ability to predict the  
38 free energy of binding means, given a larger range of substituents and attachment points, it could  
39 be possible to rank potential ligands with affinity for αvβ6 and provide guidance on the synthesis  
40 of compounds.  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Table 5: Relative free binding energy of ligand derivatives. Experimental  $\Delta\Delta G$  values have an error of 0.6 kcal mol<sup>-1</sup>.

Transformation	$\Delta\Delta G$ (kcal mol <sup>-1</sup> )		Absolute Difference
	Experimental	Calculated	
CH <sub>3</sub> → OCH <sub>3</sub>	-0.1	-2.6 ± 0.3	2.5
OCH <sub>3</sub> → OCF <sub>3</sub>	-0.3	-2.3 ± 0.3	2.0
H → F	-0.5	-1.9 ± 0.2	1.4
OCF <sub>3</sub> → CF <sub>3</sub>	-0.5	-0.5 ± 0.4	0.0
H → CH <sub>3</sub>	-1.0	-2.9 ± 0.3	1.9
CH <sub>3</sub> → CF <sub>3</sub>	-1.0	-1.4 ± 0.3	0.4
H → OCH <sub>3</sub>	-1.1	-1.8 ± 0.3	0.7
H → OCF <sub>3</sub>	-1.4	-1.9 ± 0.4	0.5
F → CF <sub>3</sub>	-1.4	-1.3 ± 0.3	0.1
H → CF <sub>3</sub>	-1.9	-2.3 ± 0.3	0.4

## Conclusions

In this work, MD and FEP simulations aid the investigation of potential antagonists of  $\alpha\text{v}\beta\text{6}$ . Consequently, CHARMM force field parameters have been developed for 1,2,3,4-tetrahydro 1,8-naphthyridine. As naphthyridine is an important moiety in medicinal chemistry,<sup>15</sup> we expect these new parameters will find utility in future computational studies. An MD simulation of  $\alpha\text{v}\beta\text{6}$  bound with the the pro domain of TGF- $\beta\text{1}$ , starting from an X-ray crystal structure,<sup>10</sup> was performed in order to understand the binding site interactions of the natural ligand. MD simulations on a series of compounds, based on an RGD mimetic, were also performed starting from docked structures. From these studies the importance of the canonical interactions, the bidentate interaction of the 1,8-naphthyridine with ( $\alpha\text{v}$ )-Asp<sup>218</sup> and the acid binding to the MIDAS site, are clear. These interactions are supported by unpublished SAR which (i) demonstrates the distance between the base and the acid is critical and (ii) shows that maintaining  $\alpha\text{v}\beta\text{6}$  affinity whilst structurally modifying the acid and 1,8-naphthyridine is particularly challenging. The calculations suggest the substituted aryl ring is essentially solvent exposed. Other work<sup>5,8,16</sup> has found that about two log units of potency can be gained with the optimal substituent both with this chemotype<sup>5,16</sup> and others,<sup>8</sup> but this involves larger substituents which can form additional interactions with the receptor.

1  
2  
3 FEP simulations have enabled us to estimate the relative free energies of binding between pairs  
4 of RGD mimetics. As the range in pIC<sub>50</sub> of the subset of compounds studied was 5.2 to 7.1, which  
5 is relatively narrow, future work should expand the substituents studied to include more potent  
6 compounds. Furthermore, in order to rank this series of β6 antagonists, all perturbations should  
7 be linked in some way so that the energies can be compared. Therefore, an alternative perturbation  
8 map should be considered. As suggested by Cournia et al.,<sup>20</sup> instead of performing perturbations to  
9 each ligand from a single reference compound, substituents should be connected in a single graph,  
10 arranged by similarity. By connecting each substituent to at least two other derivatives in a closed  
11 cycle, it is also possible to compute sampling errors as the total free energy change in a closed  
12 thermodynamic cycle should equal zero. Nevertheless, by comparing the binding free energies  
13 calculated in this study with the difference in pIC<sub>50</sub> value for each pair, we have shown that this  
14 integrin system, along with this series of ligands is amenable to study by FEP, with a good level of  
15 accuracy.  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

29 From a drug design perspective, lead optimisation of αv integrin inhibitors (αvβ6 inhibitors for  
30 example) has been driven empirically<sup>5</sup> and a priori selection of optimal substituents on the carbo-  
31 aromatic is difficult. This is because the current understanding in how the substituted aryl part of  
32 the molecule affects potency and selectivity through interactions with the specificity determining  
33 loop in the binding site is poor.<sup>3,5</sup> As a result, lead optimisation requires a large team of synthetic  
34 chemists and substantial budget and can mean progress towards a clinical candidate becomes slow.  
35 There is therefore great value in any method which transitions molecular design from empiricism  
36 to theory, which is reliable.  
37  
38  
39  
40  
41  
42  
43  
44

45 Based on our previous work,<sup>18</sup> large numbers of potential inhibitors can be generated com-  
46 putationally featuring multi-substituted aryl motifs. Whilst this has advantages, multi-substituted  
47 aromatics can be difficult to synthesise which exacerbates what are already often challenging  
48 and long syntheses of the inhibitor. So some additional method to select the best compounds  
49 to make is needed and this illustrates the value of this work as it paves the way towards a more  
50 robust computational prediction of affinity, which should be valuable in prioritising compounds for  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 synthesis.  
4  
5  
6  
7

## 8 **Conflict of Interest**

9

10 The authors declare the following competing financial interest: SJFM is a shareholder in GSK.  
11  
12  
13

## 14 **Acknowledgement**

15  
16  
17

18 E.E.G. and S.A.O. are supported by the Engineering and Physical Sciences Research Council  
19 (EPSRC) under GlaxoSmithKline CASE award schemes [EP/R512059/1 and EP/P510592/1, re-  
20 spectively]. We acknowledge the use of Athena at HPC Midlands+, which was funded by the  
21 EPSRC on grant EP/P020232/1, in this research, as part of the HPC Midlands+ consortium and we  
22 thank EPSRC for funding via Prosperity Partnership EP/S035990/1. We are grateful for access to  
23 the University of Nottingham High Performance Computer. We thank Stephen Pickett, Ian Wall,  
24 Arnaldo Filho and Tom McInally for useful discussions.  
25  
26  
27  
28  
29  
30  
31  
32  
33

## 34 **Supporting Information Available**

35  
36  
37

38 Topology and parameters developed can be found in the CHARMM format in the supporting  
39 information. PDB files containing coordinates of the docked compounds are also provided. An  
40 estimation of the computational cost of FEP simulations can also be found in the supporting  
41 information. Equilibrium MD and FEP trajectories will be uploaded to Zenodo.  
42  
43  
44  
45  
46  
47

## 48 **References**

49  
50

- 51 (1) Kotecha, A.; Wang, Q.; Dong, X.; Ilca, S. L.; Ondiviela, M.; Zihe, R.; Seago, J.; Charleston, B.;  
52 Fry, E. E.; Abrescia, N. G. A.; Springer, T. A.; Huiskonen, J. T.; Stuart, D. I. Rules of  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 engagement between  $\alpha v\beta 6$  integrin and foot-and-mouth disease virus. *Nat. Commun.* **2017**,  
4 8, 15408.  
5  
6  
7
- 8 (2) Raab-Westphal, S.; Marshall, J. F.; Goodman, S. L. Integrins as therapeutic targets: Successes  
9 and cancers. *Cancers* **2017**, 9, 1–28.  
10  
11
- 12 (3) Hatley, R. J.; Macdonald, S. J. F.; Slack, R. J.; Le, J.; Ludbrook, S. B.; Lukey, P. T. An  
13  $\alpha v$ -RGD Integrin Inhibitor Toolbox: Drug Discovery Insight, Challenges and Opportunities.  
14 *Angew. Chem., Int. Ed.* **2018**, 57, 3298–3321.  
15  
16  
17
- 18 (4) Richeldi, L.; Collard, H. R.; Jones, M. G. Idiopathic pulmonary fibrosis. *Lancet* **2017**, 389,  
19 1941–1952.  
20  
21  
22
- 23 (5) Adams, J. et al. Structure activity relationships of  $\alpha V$  integrin antagonists for pulmonary  
24 fibrosis by variation in aryl substituents. *ACS Med. Chem. Lett.* **2014**, 5, 1207–1212.  
25  
26  
27
- 28 (6) Ghatak, S.; Hascall, V. C.; Markwald, R. R.; Feghali-Bostwick, C.; Artlett, C. M.; Gooz, M.;  
29 Bogatkevich, G. S.; Atanelishvili, I.; Silver, R. M.; Wood, J.; Thannickal, V. J.; Misra, S.  
30 Transforming growth factor 1 (TGF1)-induced CD44V6-NOX4 signaling in pathogenesis of  
31 idiopathic pulmonary fibrosis. *J. Biol. Chem.* **2017**, 292, 10490–10519.  
32  
33  
34
- 35 (7) Sheppard, D. Transforming growth factor beta: a central modulator of pulmonary and airway  
36 inflammation and fibrosis. *Proc. Am. Thorac. Soc.* **2006**, 3, 413–417.  
37  
38  
39
- 40 (8) Procopiou, P. A. et al. Discovery of (S)-3-(3-(3,5-Dimethyl-1 H-pyrazol-1-yl)phenyl)-4-((R)-  
41 3-(2-(5,6,7,8-tetrahydro-1,8-naphthyridin-2-yl)ethyl)pyrrolidin-1-yl)butanoic Acid, a Non-  
42 peptidic  $\alpha v\beta 6$  Integrin Inhibitor for the Inhaled Treatment of Idiopathic Pulmonary Fibrosis.  
43 *J. Med. Chem.* **2018**, 61, 8417–8443.  
44  
45  
46  
47  
48
- 49 (9) Dong, X.; Zhao, B.; Iacob, R. E.; Zhu, J.; Koksai, A. C.; Lu, C.; Engen, J. R.; Springer, T. A.  
50 Force interacts with macromolecular structure in activation of TGF- $\beta$ . *Nature* **2017**, 542,  
51 55–59.  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 (10) Dong, X.; Hudson, N. E.; Lu, C.; Springer, T. A. Structural determinants of integrin  $\beta$ -subunit  
4 specificity for latent TGF- $\beta$ . *Nat. Struct. Mol. Biol.* **2014**, *21*, 1091–1096.  
5  
6  
7  
8 (11) Schrödinger LLC, The PyMOL Molecular Graphics System, Version 2.1.1. 2015.  
9  
10 (12) Di Leva, F. S.; Tomassi, S.; Di Maro, S.; Reichart, F.; Notni, J.; Dangi, A.; Marelli, U. K.;  
11 Brancaccio, D.; Merlino, F.; Wester, H. J.; Novellino, E.; Kessler, H.; Marinelli, L. From  
12 a Helix to a Small Cycle: Metadynamics-Inspired  $\alpha\nu\beta 6$  Integrin Selective Ligands. *Angew.*  
13 *Chem., Int. Ed.* **2018**, *57*, 14645–14649.  
14  
15  
16  
17 (13) Kapp, T. G.; Saverio Di Leva, F.; Notni, J.; Räder, A. F.; Fottner, M.; Reichart, F.; Re-  
18 ich, D.; Wurzer, A.; Steiger, K.; Novellino, E.; Kiran Marelli, U.; Wester, H.-J.; Marinelli, L.;  
19 Kessler, H. N-Methylation of isoDGR Peptides: Discovery of a Selective  $\alpha 5\beta 1$ -Integrin  
20 Ligand as a Potent Tumor Imaging Agent. *J. Med. Chem.* **2018**, *61*, 2490–2499.  
21  
22  
23 (14) Ghitti, M.; Spitaleri, A.; Valentinis, B.; Mari, S.; Asperti, C.; Traversari, C.; Rizzardi, G. P.;  
24 Musco, G. Molecular dynamics reveal that isoDGR-containing cyclopeptides are true  $\alpha\nu\beta 3$   
25 antagonists unable to promote integrin allostery and activation. *Angew. Chem., Int. Ed.* **2012**,  
26 *51*, 7702–7705.  
27  
28 (15) Madaan, A.; Verma, R.; Kumar, V.; Singh, A. T.; Jain, S. K.; Jaggi, M. 1,8-Naphthyridine  
29 Derivatives: A Review of Multiple Biological Activities. *Arch. Pharm. Chem. Life Sci.* **2015**,  
30 *348*, 837–860.  
31  
32  
33 (16) Robinson, H.; Oatley, S. A.; Rowedder, J. E.; Slade, P.; Macdonald, S. J. F.; Argent, S. P.;  
34 Hirst, J. D.; McNally, T.; Moody, C. J. Late stage functionalization via Chan-Lam amination:  
35 rapid access to potent and selective integrin inhibitors. *Chem. Eur. J. in press*.  
36  
37 (17) Barrett, T. N. et al. Profile of a Highly Selective Quaternized Pyrrolidine Betaine  $\alpha\nu\beta 6$  Integrin  
38 Inhibitor—(3S)-3-(3-(3,5-Dimethyl-1H-pyrazol-1-yl)phenyl)-4-((1S and 1R,3R)-1-methyl-3-  
39 (2-(5,6,7,8-tetrahydro-1,8-naphthyridin-2-yl)ethyl)pyrrolidin-1-ium-1-yl)butanoate Synthe-  
40 siz. *J. Med. Chem.* **2019**, *62*, 7543–7556.  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 (18) Oglic, D.; Oatley, S. A.; Macdonald, S. J. F.; McNally, T.; Garnett, R.; Hirst, J. D.; Gärtner, T.  
4 Active Search for Computer-aided Drug Design. *Mol. Inf.* **2018**, *37*, 1–15.  
5  
6  
7  
8 (19) Bash, P. A.; Chandra Singh, U.; Brown, F. K.; Langridge, R.; Kollman, P. A. Calculation of  
9 the relative change in binding free energy of a protein-inhibitor complex. *Science* **1987**, *235*,  
10 574–576.  
11  
12  
13  
14 (20) Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug  
15 Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017**, *57*,  
16 2911–2937.  
17  
18  
19  
20  
21 (21) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; De Groot, B. L.; Grubmüller, H.;  
22 MacKerell, A. D. CHARMM36m: An improved force field for folded and intrinsically disor-  
23 dered proteins. *Nat. Methods* **2016**, *14*, 71–73.  
24  
25  
26  
27  
28 (22) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.;  
29 Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A. D. CHARMM general force field: A  
30 force field for drug-like molecules compatible with the CHARMM all-atom additive biological  
31 force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.  
32  
33  
34  
35  
36  
37 (23) Vanommeslaeghe, K.; MacKerell, A. D. Automation of the CHARMM General Force Field  
38 (CGenFF) I: Bond Perception and Atom Typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144–3154.  
39  
40  
41  
42 (24) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General  
43 Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J.*  
44 *Chem. Inf. Model.* **2012**, *52*, 3155–3168.  
45  
46  
47  
48  
49 (25) Shao, Y. et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program  
50 package. *Mol. Phys.* **2015**, *113*, 184–215.  
51  
52  
53  
54 (26) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison  
55 of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.  
56  
57  
58

- 1  
2  
3 (27) Guvench, O.; MacKerell, A. D. Automated conformational energy fitting for force-field de-  
4 velopment. *Journal of Molecular Modeling* **2008**, *14*, 667–679.  
5  
6  
7  
8 (28) McGann, M. FRED pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.*  
9 **2011**, *51*, 578–596.  
10  
11  
12 (29) Hawkins, P. C.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T. Conformer  
13 generation with OMEGA: Algorithm and validation using high quality structures from the  
14 protein databank and cambridge structural database. *J. Chem. Inf. Model.* **2010**, *50*, 572–584.  
15  
16  
17 (30) Halgren, T. A. MMFF VI. MMFF94s option for energy minimization studies. *J. Comput.*  
18 *Chem.* **1999**, *20*, 720–729.  
19  
20  
21 (31) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A web-based graphical user interface for  
22 CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865.  
23  
24  
25 (32) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D.  
26 Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved  
27 Sampling of the Backbone,  $\psi$  and Side-Chain  $\chi_1$  and  $\chi_2$  Dihedral Angles. *J. Chem. Theory*  
28 *Comput.* **2012**, *8*, 3257–3273.  
29  
30  
31 (33) Beglov, D.; Roux, B. Finite representation of an infinite bulk system: Solvent boundary  
32 potential for computer simulations. *J. Chem. Phys.* **1994**, *100*, 9050–9063.  
33  
34  
35 (34) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.;  
36 Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput.*  
37 *Chem.* **2005**, *26*, 1781–802.  
38  
39  
40 (35) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics  
41 simulation: The Langevin piston method. *J. Comput. Phys.* **1995**, *103*, 4613–4621.  
42  
43  
44 (36) Feller, S. E.; Pastor, R. W.; Rojnuckarin, A.; Bogusz, S.; Brooks, B. R. Effect of Electrostatic  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- Force Truncation on Interfacial and Transport Properties of Water. *J. Phys. Chem.* **1996**, *100*, 17011–17020.
- (37) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (38) Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M. Hidden thermodynamics of mutant proteins. *Science* **1989**, *244*, 1069–1072.
- (39) Pearlman, D. A. A comparison of alternative approaches to free energy calculations. *J. Phys. Chem.* **1994**, *98*, 1487–1493.
- (40) Michel, J.; Essex, J. W. Prediction of protein–ligand binding affinity by free energy simulations: assumptions, pitfalls and expectations. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 639–658.
- (41) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J. Chem. Phys.* **1994**, *100*, 9025–9031.
- (42) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett.* **1994**, *222*, 529–539.
- (43) Kiang, Y.-H.; Xu, W.; Stephens, P. W.; Ball, R. G.; Yasuda, N. Layered Structure and Swelling Behavior of a Multiple Hydrate-Forming Pharmaceutical Compound. *Cryst. Growth Des.* **2009**, *9*, 1833–1843.
- (44) Chen, X.; Zhao, H.; Chen, C.; Jiang, H.; Zhang, M. Transfer hydrogenative para -selective aminoalkylation of aniline derivatives with N-heteroarenes via ruthenium/acid dual catalysis. *Chem. Commun. (London)* **2018**, *54*, 9087–9090.

- 1  
2  
3 (45) Ma, W.; Chen, F.; Liu, Y.; He, Y.-M.; Fan, Q.-H. Ruthenium-Catalyzed Enantioselective  
4 Hydrogenation of 1,8-Naphthyridine Derivatives. *Org. Lett.* **2016**, *18*, 2730–2733.  
5  
6  
7  
8 (46) A. Kinney, W.; K. Luci, D.; J. Santulli, R.; A. Gauthier, D.; A. Tounge, B.; Ghosh, S.; C.  
9 Proost, J.; De Corte, B.; A. Galemme, R.; M. Lewis, J.; E. Dorsch, W.; W. Wagaman, M.;  
10 P. Damiano, B.; E. Maryanoff, B. A Concise Synthesis of an Indenopyrrolidine-based Dual  
11 avb3/avb5 Integrin Antagonist. *Heterocycles* **2004**, *62*, 543.  
12  
13  
14  
15  
16 (47) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database.  
17 *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2016**, *72*, 171–179.  
18  
19  
20  
21 (48) Xiong, J. P.; Stehle, T.; Zhang, R.; Joachimiak, A.; Frech, M.; Goodman, S. L.; Arnaout, M. A.  
22 Crystal structure of the extracellular segment of integrin  $\alpha V\beta 3$  in complex with an Arg-Gly-  
23 Asp ligand. *Science* **2002**, *296*, 151–155.  
24  
25  
26  
27 (49) Xia, W.; Springer, T. A. Metal ion and ligand binding of Integrin  $\alpha 5\beta 1$ . *Proc. Natl. Acad. Sci.*  
28 *U. S. A.* **2014**, *111*, 17863–17868.  
29  
30  
31  
32 (50) Maltsev, O. V.; Marelli, U. K.; Kapp, T. G.; Di Leva, F. S.; Di Maro, S.; Nieberler, M.;  
33 Reuning, U.; Schwaiger, M.; Novellino, E.; Marinelli, L.; Kessler, H. Stable Peptides Instead  
34 of Stapled Peptides: Highly Potent  $\alpha v\beta 6$ -Selective Integrin Ligands. *Angew. Chem., Int. Ed.*  
35 **2016**, *55*, 1535–1539.  
36  
37  
38  
39 (51) Civera, M.; Arosio, D.; Bonato, F.; Manzoni, L.; Pignataro, L.; Zanella, S.; Gennari, C.;  
40 Piarulli, U.; Belvisi, L. Investigating the interaction of cyclic RGD peptidomimetics with  
41  $\alpha V\beta 6$  integrin by biochemical and molecular docking studies. *Cancers* **2017**, *9*, 1–13.  
42  
43  
44  
45 (52) Gallivan, J. P.; Dougherty, D. A. Cation- $\pi$  interactions in structural biology. *Proc. Natl. Acad.*  
46 *Sci. U. S. A.* **1999**, *96*, 9459–9464.  
47  
48  
49  
50 (53) Kumar, K.; Woo, S. M.; Siu, T.; Cortopassi, W. A.; Duarte, F.; Paton, R. S. Cation- $\pi$   
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 interactions in protein-ligand binding: Theory and data-mining reveal different roles for  
4 lysine and arginine. *Chem. Sci.* **2018**, *9*, 2655–2665.  
5  
6  
7
- 8 (54) Duffy, E. M.; Jorgensen, W. L.; Kowalczyk, P. J. Do Denaturants Interact with Aromatic  
9 Hydrocarbons in Water? *J. Am. Chem. Soc.* **1993**, *115*, 9271–9275.  
10  
11  
12
- 13 (55) Tsuzuki, S.; Mikami, M.; Yamada, S. Origin of attraction, magnitude, and directionality of  
14 interactions in benzene complexes with pyridinium cations. *J. Am. Chem. Soc.* **2007**, *129*,  
15 8656–8662.  
16  
17  
18
- 19 (56) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. Guidelines for the analysis of free energy  
20 calculations. *J. Comput.-Aided Mol. Des.* **2015**, *29*, 397–411.  
21  
22  
23
- 24 (57) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J.*  
25 *Comput. Phys.* **1976**, *22*, 245–268.  
26  
27  
28
- 29 (58) Liu, P.; Dehez, F.; Cai, W.; Chipot, C. A Toolkit for the Analysis of Free-Energy Perturbation  
30 Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 2606–2616.  
31  
32  
33
- 34 (59) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Protein-ligand docking: Current status and future  
35 challenges. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 15–26.  
36  
37  
38
- 39 (60) Wang, L. et al. Accurate and reliable prediction of relative ligand binding potency in prospec-  
40 tive drug discovery by way of a modern free-energy calculation protocol and force field. *J.*  
41 *Am. Chem. Soc.* **2015**, *137*, 2695–2703.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60