Gavin Morris, Stoyan Stoychev, Previn Naicker, Heini W. Dirr and Sylvia Fanucchi*

The forkhead domain hinge-loop plays a pivotal role in DNA binding and transcriptional activity of FOXP2

https://doi.org/10.1515/hsz-2018-0185
Received March 15, 2018; accepted April 26, 2018

Abstract: Forkhead box (FOX) proteins are a ubiquitously expressed family of transcription factors that regulate the development and differentiation of a wide range of tissues in animals. The FOXP subfamily members are the only known FOX proteins capable of forming domain-swapped forkhead domain (FHD) dimers. This is proposed to be due to an evolutionary mutation (P539A) that lies in the FHD hinge loop, a key region thought to fine-tune DNA sequence specificity in the FOX transcription factors. Considering the importance of the hinge loop in both the dimerisation mechanism of the FOXP FHD and its role in tuning DNA binding, a detailed investigation into the implications of mutations within this region could provide important insight into the evolution of the FOX family. Isothermal titration calorimetry and hydrogen exchange mass spectroscopy were used to study the thermodynamic binding signature and changes in backbone dynamics of FOXP2 FHD DNA binding. Dual luciferase reporter assays were performed to study the effect that the hinge-loop mutation has on FOXP2 transcriptional activity *in vivo*. We demonstrate that the change in dynamics of the hinge-loop region of FOXP2 alters the energetics and mechanism of DNA binding highlighting the critical role of hinge loop mutations in regulating DNA binding characteristics of the FOX proteins.

Keywords: backbone dynamics; DNA-binding; electrostatics; forkhead; FOXP; hinge-loop.

***Corresponding author: Sylvia Fanucchi,** Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of the Witwatersrand, 1 Jan Smuts Ave, Braamfontein, 2050 Johannesburg, Gauteng, South Africa,
e-mail: sylvia.fanucchi@wits.ac.za
Gavin Morris and Heini W. Dirr: Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of the Witwatersrand, 1 Jan Smuts Ave, Braamfontein, 2050 Johannesburg, Gauteng, South Africa
Stoyan Stoychev and Previn Naicker: CSIR Biosciences, CSIR, Meiring Naude Road, Brummeria, 0001 Pretoria, Gauteng, South Africa

Introduction

The forkhead box (FOX) family is a prominent transcription factor family responsible for the differentiation and growth regulation of a large variety of cell types across multiple vertebrate taxa. FOX transcription factors play a role in numerous processes including development and differentiation, cell cycle regulation and homeostasis (Hannenhalli and Kaestner, 2009).

The FOX family of transcription factors is identified by a highly conserved DNA binding domain, the forkhead domain (FHD) which is a variant of a helix-turn-helix motif termed a winged helix (Clark et al., 1993; Gajiwala and Burley, 2000). FOX family members are subdivided into 19 (A–S) subfamilies based on sequence divergence from the FHD and homology of the flanking sequences (Lai et al., 1993; Kaestner et al., 2000). All the FOX proteins are multidomain proteins and show substantial deviation in the number and type of domains flanking the FHD. The FOXP subfamily (FOXP1-4) is characterised by a leucine zipper and zinc finger domain located N-terminal to the FHD. FOXP1, 2 and 4 also possess a long N-terminal polyglutamine tract of unknown function (Lai et al., 2001).

FOXP2, the first gene product to be associated with language acquisition and use, plays a vital part in the development of brain regions necessary for complex language comprehension (Lai et al., 2001). Loss-of-function mutations in the FOXP2 FHD cause a congenital mental retardation that presents as a difficulty in fine motor control of the orofacial muscles required for complex language production (orofacial dyspraxia) as well as reduced language cognition (Fujita et al., 2008). FOXP2 also plays a significant role in non-communicable disease. Aberrant expression of FOXP2 has been observed in breast cancer, diffuse large B-cell lymphoma and multiple myeloma (Cuiffo et al., 2014; Wong et al., 2016). FOXP2 has also been implicated in schizophrenia, autism and most recently Huntington's disease (Gong et al., 2004; Spaniel et al., 2011; Hachigian et al., 2017).

Three of the four FOXP subfamily members (FOXP1/2/4) share overlapping regions of expression in the developing and adult vertebrate brain (Teramitsu, 2004;

Takahashi et al., 2009). As a result of this overlapping expression pattern as well as the presence of a conserved leucine zipper domain, the FOXP subfamily is unique among FOX transcription factors in its ability to form both homo- and heterotypic dimers via the leucine zipper domain (Li et al., 2004). Dimerisation of the FOXP subfamily members has been shown to be essential *in vivo* and is responsible for the differential expression of a large array of target genes (Li et al., 2004; Sin et al., 2014). In addition to dimerisation through the leucine zipper, crystal structures reveal that the isolated FOXP FHD is capable of forming three-dimensional (3D)-domain swapped dimers (Stroud et al., 2006). Domain swapping is an uncommon form of oligomerisation whereby a portion of the associating domains is exchanged with the equivalent region in the oligomerisation partner. This is usually facilitated through a 'hinge-loop' – a random coil between the exchanged and fixed regions of the monomer (Rousseau et al., 2003). The FOXP FHD has a far greater propensity to form a domain swapped dimer than the FHD of other members of the FOX family. This is attributed to a single proline to alanine substitution in the hinge-loop between helix 2 and 3 that promotes extension of helix 2 into the hinge-loop (Stroud et al., 2006). Indeed, reversal of this hinge-loop mutation successfully prevents dimerisation of the FOXP FHD (Stroud et al., 2006; Chu et al., 2011).

All FOX FHD structures reveal that the FHD interacts with DNA via helix 3 – the recognition helix – which inserts into the major groove (Clark et al., 1993; Stroud et al., 2006). Despite almost complete sequence conservation within the recognition helix of the FHD as well as high conservation of the recognition DNA sequence (TA/GTTT/GG/AA/GT/C), the FOX family members display a large deviation in target gene activity (Pierrou et al., 1994). Furthermore the FOX FHD has been shown to bind to more than one DNA sequence with varying rates and affinity and sequence specificity is believed to be a form of transcriptional regulation (Webb et al., 2017). The region of the FOX FHD that is considered most likely to contribute to sequence specificity is the region immediately prior to the recognition helix consisting of helix 2 and the less highly conserved hinge-loop region (Figure 1) (Overdier et al., 1994; Pierrou et al., 1994; Marsden et al., 1998).

It is interesting, from both an evolutionary and functional perspective that of all the FOX proteins, (of which there are over 100) only the four members of the FOXP subfamily have been discovered to have developed the ability to form FHD dimers. This unique ability is entirely due to the amino acid sequence within the hinge-loop region and has lead us to question the significance of this region in FOX protein function. Indeed, mutations in the hinge-loop of FOXP3 that have been shown to reduce

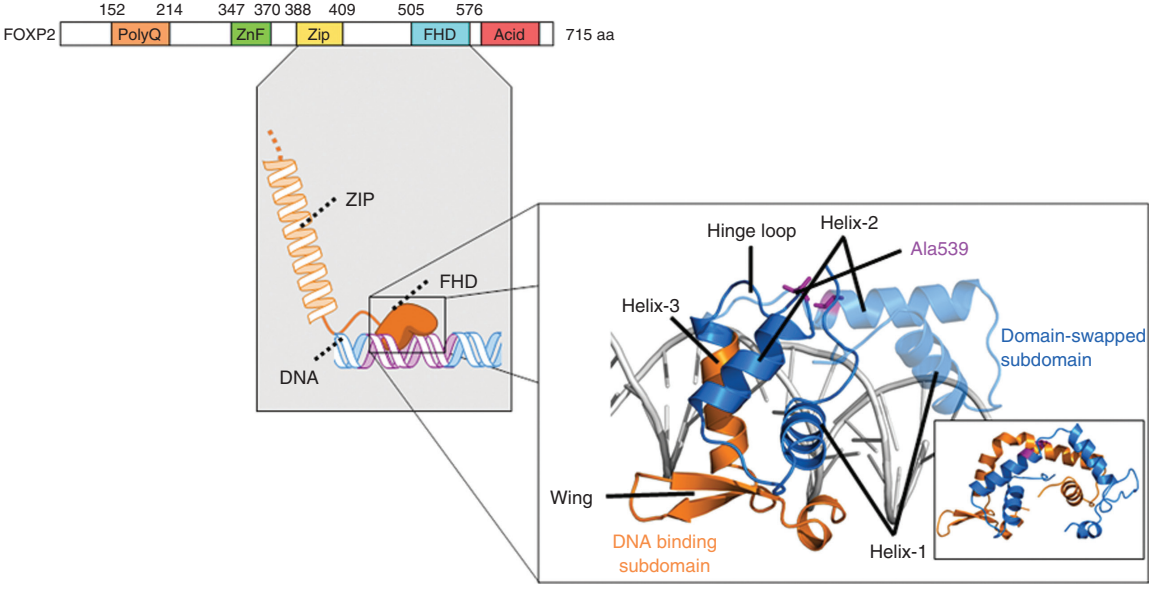

Figure 1: Anatomy of the FOXP2 forkhead domain.
FOXP2 has two dimerisation domains: the leucine zipper (ZIP) and the forkhead domain (FHD). The FOXP2 FHD can form a domain-swapped dimer (inset) whereby associated FOXP2 FHD monomers exchange the domain-swapped subdomain (helix-1 and -2; blue). The exchange (transparent) involves the extension of helix-2 into the hinge loop that connects the two subdomains of the FOXP2 FHD. This is thought to be possible because of an evolutionary hinge loop mutation (P539A) exclusive to the FOXP subfamily (Ala539 shown in purple). The DNA binding subdomain (helix-3 and the wing; orange) are responsible for the formation of the protein-DNA interface. The figure was produced using the crystal structure of the FOXP2 FHD bound to DNA (PDB ID: 2A07).

dimer propensity have also been associated with disease (Bandukwala et al., 2011). Does this mean that the formation of FHD dimers specifically (over and above dimers formed via the leucine zipper) is significant for FOXP function? In this work we aim to determine the role that the FOXP proline to alanine hinge-loop mutation has on the DNA binding mechanism of the FOXP2 FHD. This study will provide insight into the complexities of transcriptional regulation via FOXP proteins and will distinguish between proteins with an exclusively monomeric FHD and those with the potential to form domain swapped dimers, hopefully providing a clue as to whether the FOXP subfamily has evolved to develop a unique mode of action compared to the other members of the family.

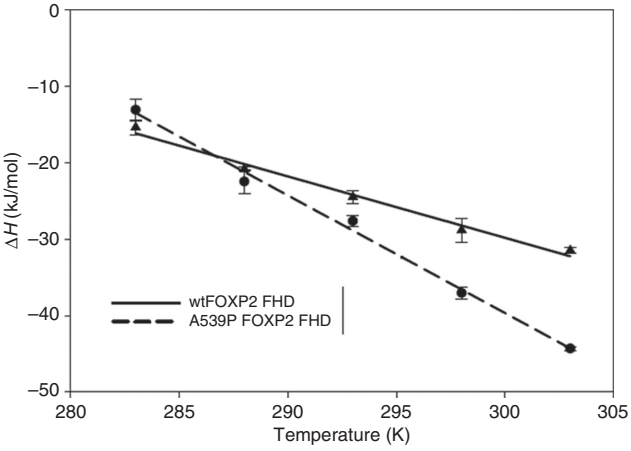

Figure 2: Temperature dependence of DNA binding of the wild-type (triangles and solid line) and A539P (circles and dashed line) FOXP2 FHD.
The heat capacity values are $\Delta C_p = -0.88 \pm 0.048$ kJ/mol/K for wild-type and $\Delta C_p = -1.56 \pm 0.04$ kJ/mol/K for the A539P FOXP2 FHD. Error bars are the sample standard deviation of the average of at least duplicate or triplicate titrations.

Results

In order to compare the DNA binding mechanism of a FOX FHD that has the potential to form a dimer to one that does not, we have worked with both the wild-type FOXP2 FHD and the exclusively monomeric A539P hinge-loop mutant. The thermodynamics of DNA binding has been investigated in this study at different temperatures and at different salt concentrations to obtain information on the heat capacity and the electrostatics of binding, respectively. The dynamics and flexibility of the wild-type and A539P mutant upon DNA binding has been studied so as to assess whether restrictions on backbone flexibility affect DNA binding. Finally, a luciferase reporter assay was also performed so as to understand the effect of the monomeric mutation on FOXP2 transcriptional ability and hence to provide physiological relevance to the *in vitro* work.

Heat capacity of DNA binding

The heat capacity of biochemical interactions provides insight into the arrangements and types of bonds involved in the interaction. When DNA binds to a protein, a number of factors will contribute to the change in heat capacity (ΔC_p) upon DNA binding. These factors include the structural changes that may accompany DNA binding, changes in ionic shielding as well as changes in the solvation of the protein and DNA. Usually, sequence-specific DNA binding is accompanied by a large negative heat capacity change, while non sequence specific binding does not necessarily yield a large negative value (Prabhu and Sharp, 2005).

Temperature-dependant DNA binding studies of the wild-type FOXP2 FHD and the A539P mutant show significant differences (Figure 2). The wild-type FOXP2 FHD has an unusually small ΔC_p upon binding DNA of -0.88 ± 0.05 kJ/mol/K. This is nearly half of that obtained for the A539P FOXP2 FHD mutant where $\Delta C_p = -1.56 \pm 0.04$ kJ/mol/K. This suggests that the binding of the wild-type is less sequence specific than the A539P mutant and may involve more backbone rather than base interactions. The data for both the wild-type and A539P FOXP2 FHD agree well with the linear model fitted for determining heat capacity indicating that no additional temperature-dependent equilibria with a significant enthalpy change occur across the temperature range of 10–30°C.

Salt studies of binding

The polyelectrolytic nature of DNA results in the formation of steep counter-ion concentration gradients around the phosphate backbone, a process known as counter-ion condensation (deHaseth et al., 1977; Manning, 1978; Record et al., 1978; Privalov et al., 2011). Exclusion of counter-ions from the DNA backbone during protein binding to the DNA results in a measurable cratic entropy of mixing that can be used to dissect the total entropic term into salt-dependent (electrostatic) and salt-independent (non-electrostatic) components. This analysis can be performed by fitting a linear regression to the double log plot of the observed association constant ($\log K_a$) as a function of salt concentration (log[S]) (Figure 3A) (Privalov et al., 2011).

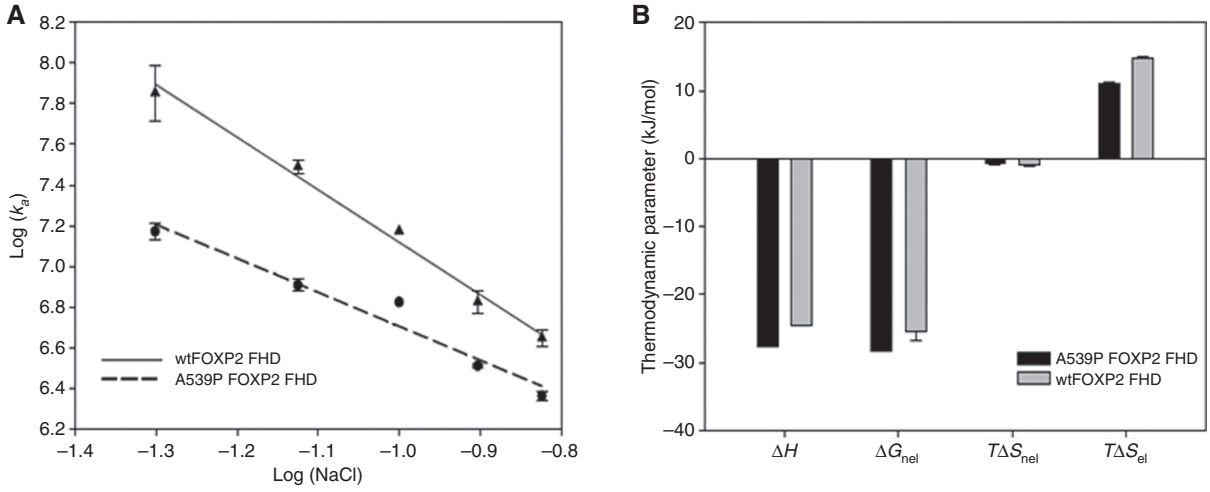

Figure 3: Salt dependent studies of DNA binding by the wild-type and A539P FOXP2 FHD.
(A) The number of counter-ions displaced (*N*) during binding can be determined from the gradient of the linear regression which can then be related to the number of ionic contacts made. The steeper gradient for the wild-type (solid line) implies a greater number of backbone interactions with the DNA compared to the mutant (dashed line). Error bars are the sample standard deviation of the average of at least duplicate or triplicate titrations. The ΔH displays a minor dependence on the salt concentrations used in this study (Supplementary Figure 6). (B) The entropy of binding at 100 mM NaCl was dissected into electrostatic ($T\Delta S_{el}$) and non-electrostatic ($T\Delta S_{nel}$) components using the enthalpy of binding (ΔH) averages from all salt concentrations. The wild-type (grey) displays a greater entropy of binding but the magnitude of the enthalpy and non-electrostatic free energy of binding is smaller compared to the mutant (black).

The affinity constant at 1 M salt (K_{nel}) is presumed to be a measure of the salt independent (non-electrostatic) component of the free energy of binding (Privalov et al., 2011). The total entropy change upon binding can then be dissected into electrostatic (ΔS_{el}) and non-electrostatic (ΔS_{nel}) components. The number of counter-ions (*N*) excluded from the DNA backbone during formation of the protein-DNA interface can be obtained from the gradient of the linear regression by considering the DNA phosphate counter-ion occupancy of NaCl ($\psi = 0.64$) (Olmsted et al., 1995; Privalov et al., 2011).

As with the heat capacities, there is a difference in the DNA binding electrostatics of the wild-type FOXP2 FHD compared to the A539P mutant. The wild-type displays a significant dependence on electrostatic interactions with most of the entropy originating from the electrostatic component ($T\Delta S_{el} = 14.64 \pm 0.38$ kJ/mol) while the non-electrostatic contribution to the entropy, ($T\Delta S_{nel}$) is -1 ± 0.38 kJ/mol (Figure 3B). The A539P mutant FOXP2 FHD has a smaller $T\Delta S_{el}$ component than the wild-type with a value of 7.53 ± 0.16 kJ/mol and a lower $T\Delta S_{nel}$ of -4.11 ± 0.16 kJ/mol (Figure 3B). As would be expected following these results, the wild-type FOXP2 FHD makes a higher number of ionic contacts with the DNA backbone than the does the A539P mutant. The wild-type forms approximately four ($N = 2.58 \pm 0.27$) ionic contacts with the DNA while the mutant forms approximately two to three ($N = 1.67 \pm 0.028$).

Dynamics of DNA binding

As differences in the thermodynamic DNA binding signatures were observed between the wild-type FOXP2 FHD and the A539P mutant, and particularly as a difference in the conformational entropy component was found, hydrogen-deuterium exchange experiments were performed in the absence and presence of DNA so as to identify the regions of the FHD responsible for the observed differences (Figure 4).

The wild-type FOXP2 FHD displays a decrease in backbone solvent accessibility upon binding to DNA in regions presumed to be shielded by the protein-DNA interface (helix-3 and wing-1) as well as, somewhat surprisingly, in regions distal from the interface (helix-1 and helix-2). Interestingly, the A539P FOXP2 FHD resembles the less dynamic, less solvent exposed DNA-bound form of the wild-type whether it is in the presence or absence of DNA suggesting that the flexibility of the hinge-loop plays a significant role in the global dynamics of the FHD.

As the wild-type showed a change in backbone dynamics upon DNA binding, fluorescence quenching studies were performed in the presence of a collisional quencher, in order to assess whether the changes in backbone dynamics are as a result of significant refolding events. Stern-Volmer constants were determined for both the wild-type and A539P FOXP2 FHD in the absence and presence of DNA (Figure 5). Despite only the wild-type showing a change in backbone dynamics upon

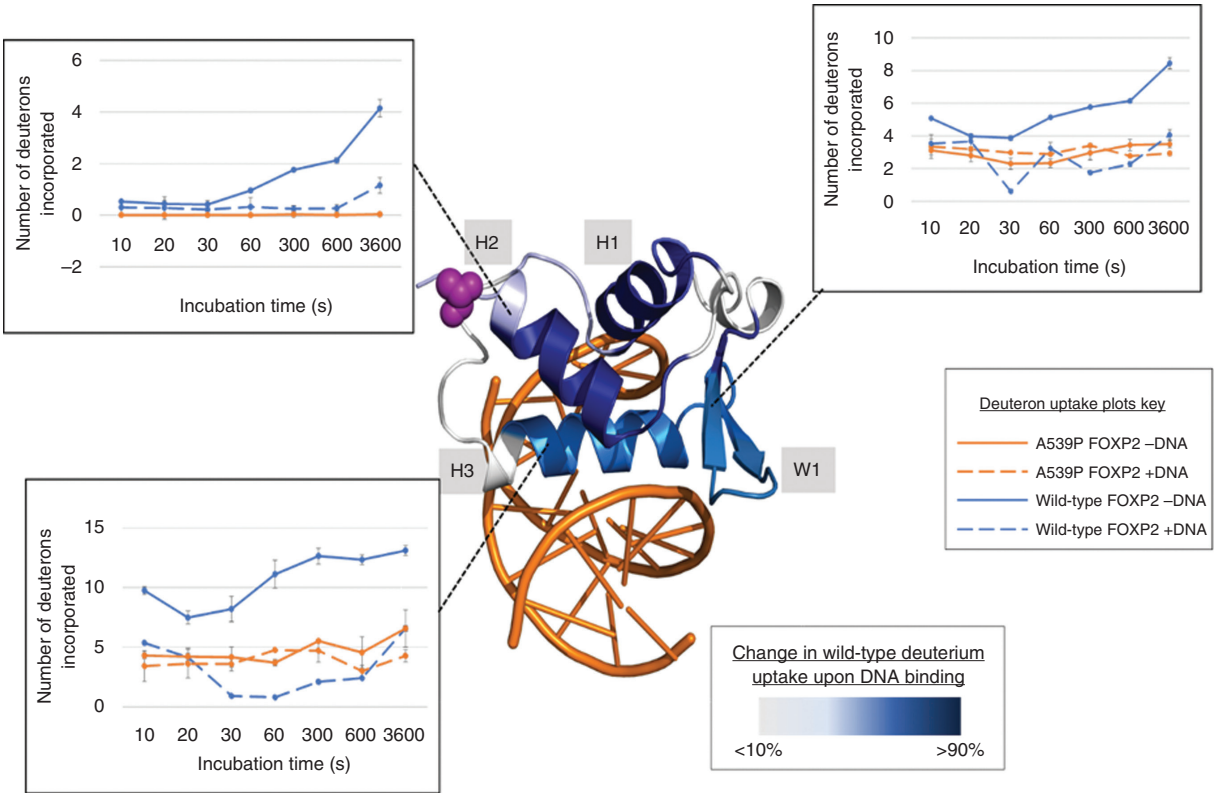

Figure 4: The change in backbone dynamics of the wild-type FOXP2 FHD induced by DNA binding as observed with hydrogen-deuterium exchange mass spectrometry.
The change in deuterium uptake is calculated by subtracting deuterium uptake for the DNA bound wild-type FOXP2 FHD from the deuterium uptake from the unbound form at the incubation time of 3600 s. The difference in deuteron uptake is mapped to the crystal structure of the monomeric wild-type FOXP2 FHD bound to the consensus DNA sequence (PDB ID: 2A07). The site of Ala539 in the hinge-loop is shown in purple. Insets: Deuterium uptake plots for helix-2 (H2; Ser522-Thr535), the recognition helix (H3; Trp548-His559) and the β-sheet wing (W1; Cys561-Glu577) of the wild-type (blue) and A539P (orange) FOXP2 FHD in the presence (dashed line) and absence (solid line) of DNA.

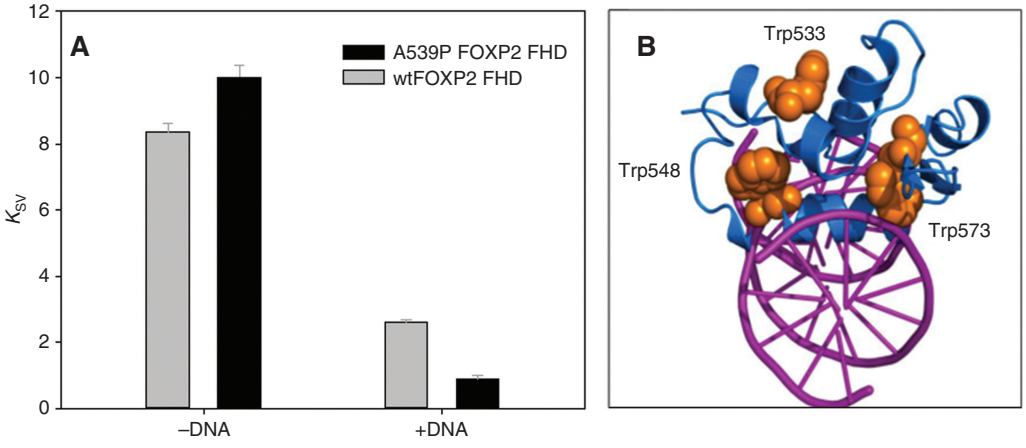

Figure 5: Collisional quenching of tryptophan fluorescence.
(A) Stern-Volmer constants (K_{sv}) for the free and DNA bound forms of the wild-type (grey) and A539P (black) FOXP2 FHD. There is a substantial reduction in tryptophan fluorescence quenching upon DNA binding of both wild-type and mutant, however, the wild-type displays greater quenching than the mutant. Hence wild-type tryptophan residues are more exposed upon DNA binding than the mutant. Errors represent the sample standard deviation of three independent replicates. (B) The positions of the three tryptophan residues in the FOXP2 FHD.

binding, both proteins showed a significant reduction in the Stern-Volmer constants upon binding to DNA implying a change in the local environment of the tryptophan residues. The FOXP2 FHD contains three tryptophan residues; two flanking the hinge loop (Trp533 and Trp548) and a third within the loop connecting the recognition helix and the β-sheet wing (Trp573). The largest contributor to the observed decrease in the Stern-Volmer constant upon binding is likely Trp573 which becomes entrenched in the protein-DNA interface. The higher solvent accessibility of the tryptophan residues in the DNA bound form of the wild-type FOXP2 FHD compared to the mutant corroborates the findings of the deuterium exchange experiments suggesting that the wild-type binds in a more open conformation.

Luciferase reporter assay

The FOXP subfamily members are the only FOX proteins known to form homo- and heterotypic dimers. FOXP dimerisation can occur at two interfaces, i.e. via the leucine zipper domain and also via the FHD. Dimerisation through the leucine zipper domain has been shown to be necessary for correct FOXP2 function *in vivo* (Li et al., 2004). However, no studies have attempted to understand the role of only the FOXP2 FHD dimerisation *in vivo*. Indeed, as the domain swapped dimer has only been identified in crystal structures of the isolated FHD, there is question as to whether dimerisation of the FHD is physiologically significant or whether it is simply an artefact of the crystallisation conditions.

Luciferase reporter assays were performed with the full length proteins that therefore contain both the FHD and the leucine zipper domain. The assays were performed with both wild-type and A539P mutant FOXP2 to assess whether the dimerisation of the FHD has any detectable role in transcriptional activity (Figure 6). Overexpression of wild-type and A539P mutant FOXP2 resulted in an approximately seven-fold and five-fold induction of luciferase activity, respectively. Thus disruption of FHD dimerisation via the A539P mutation significantly lowers transcriptional activity of FOXP2 but does not inactivate it (**$p < 0.012$).

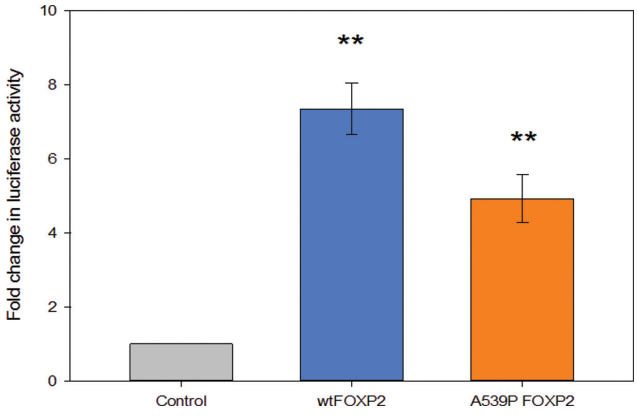

Figure 6: Dual luciferase reporter assay for FOXP2 in HEK293 cells. Reporter assays were performed with a vector encoding firefly luciferase under the control of a promoter containing six sequential cognate binding sites (TTAGG**TGTTTAC**TTTCA). Cells were transfected with an empty pcDNA4 vector (grey) or pcDNA4 vectors designed to transiently over express either wild-type FOXP2 (blue) or A539P mutant (orange). Error bars represent the standard deviation of at least three independent replicates.

Discussion

The FOXP subfamily is unique among FOX proteins in its ability to form dimers. Particularly it has the capacity to form dimers at two interfaces: the leucine zipper and the FHD. The focus of this work has been on the FHD dimer. FOXP proteins are capable of forming FHD dimers due to a proline to alanine substitution in the hinge-loop region of the FHD that allows domain swapping to occur. Here we study the effect that this hinge-loop mutation has on the DNA binding mechanism of the protein in an attempt to obtain evidence that the FOXP subfamily has evolved a novel mechanism for DNA binding, recognition, specificity and hence transcriptional regulation. We have accomplished this by isolating the wild-type FOXP2 FHD and the A539P FOXP2 FHD mutant and comparing their interaction with DNA.

Changes induced upon DNA binding

In our previous work, we examined the structure and DNA-binding thermodynamics of the wild-type FOXP2 FHD and the A539P mutant (Morris and Fanucchi, 2016). We found that there is little structural difference between the wild-type and the A539P mutant. In fact, we only detected a low proportion of wild-type dimer, concluding that the apo wild-type existed almost entirely as a monomer under the conditions we used for the study (Morris and Fanucchi, 2016). Despite this, we found that there was a distinct difference between the wild-type and the A539P mutant in their thermodynamic DNA binding signatures (Morris and Fanucchi, 2016). In this work, we expand on this discovery and clearly establish that the source of this difference lies in the role of the hinge-loop in regulating DNA binding dynamics.

The hydrogen-deuterium exchange experiments (Figure 4) suggest that the wild-type FOXP2 FHD has a more flexible and solvent exposed backbone in the absence of DNA than the A539P mutant. There is a distinct change in wild-type backbone solvent accessibility upon binding DNA where the backbone becomes less solvent exposed. The A539P mutant, on the other hand, shows little change in solvent accessibility whether in the presence or absence of DNA and, in fact, in both instances it resembles the more rigid conformation of the DNA-bound wild-type. This is an interesting result because it implies that a reduction in the flexibility of the hinge-loop as introduced by the mutation causes a global reduction in the dynamics of the entire protein. It makes sense that a weak interaction network exists between the two subdomains that are exchanged during domain swapping of the FOXP2 FHD. The hinge-loop appears to control this network. Thus the reduction in backbone flexibility caused by the A539P mutation in the hinge-loop may prevent rearrangement of secondary and tertiary structural interactions that occur freely in the wild-type FOXP2 FHD. When the more dynamic wild-type binds to DNA, its backbone becomes more restricted and resembles that of the A539P mutant. However, the fluorescence quenching results indicate that the wild-type tryptophan residues remain more exposed to the solvent when in the presence of DNA than the mutants tryptophan residues (Figure 4). Considering that the location of Trp537 indicates that it is likely to be a direct reporter of events at the protein-DNA interface, the quenching results imply that the wild-type FHD does not insert as deeply into the DNA major groove as the mutant. This could be because the inherent flexibility in the wild-type FHD might reduce the amount of time it remains associated with the DNA, preventing it from inserting as deeply into the major groove as the more conformationally restricted A539P mutant does. Furthermore, Trp548 (which lies proximal to the hinge loop) is also likely to be a significant contributor to the difference in signal observed in the dynamic fluorescence quenching study of the wild-type and mutant FOXP2 FHD (Figure 5). The hinge loop of the FHD has been shown to exist in either a random coil or α-helical conformation depending on the subfamily under question (Marsden et al., 1998; Chu et al., 2011). The different conformations of the hinge loop are thought to govern the orientation of the recognition helix in the major groove of DNA, fine tuning specific DNA contacts between the recognition helix and the DNA bases (Marsden et al., 1998). The exchange of an alanine residue with proline in the hinge loop of FOXP2 could increase the solvent accessibility of Trp548 by restricting the dynamics of the hinge loop to a more open conformation. Considering the marked difference in solvent accessibility between the wild-type and mutant FOXP2 FHD observed in this study, it appears that the residues in the C-terminal portion of the hinge loop, particularly Ala539, have a significant role in determining the structural conformation of this important region and in doing so fine tune the DNA contacts made during sequence specific binding. Such an observation is corroborated by the large change in thermodynamic binding signatures seen in this study (Figures 1 and 2). The much larger reduction in tryptophan solvent accessibility observed in the A539P mutant, compared to the wild-type upon DNA binding, suggests that a conformational switch to a more concealed conformation occurs in the mutant hinge loop not seen in the wild-type. Such a conformational switch undoubtedly contributes to the observed difference in the ΔC_p of binding and provides some evidence that this residue is a key component in defining the DNA binding of the FHD, perhaps by acting as a small energetic checkpoint for sequence recognition and specificity.

Interactions formed with the DNA

It is largely accepted that the heat capacity derived from macromolecular interactions is due to the hydration/dehydration of polar and non-polar residues. This has been shown to be true for protein-DNA interactions, although it must be noted that other contributing factors have also been suggested including binding coupled refolding of disordered regions (Murphy and Freire, 1992; Spolar and Record, 1994; Liu et al., 2008). The negative heat capacity values upon DNA binding of both wild-type and mutant (Figure 2) indicate solvation of polar groups and burial of hydrophobic groups at the interface. However, the larger value of ΔC_p obtained for the A539P mutant suggests that its DNA-protein interface could be larger possibly owing to a different orientation of the recognition helix in the DNA major groove (Parbhu and Sharp, 2005). The less negative value for the wild-type suggests there are more polar residues (and hence more water molecules) at the interface of the wild-type with DNA which results in less freedom of movement and a lower capacity to retain heat (Parbhu and Sharp, 2005). The idea that the wild-type-DNA interface is more polar is corroborated by the salt study (Figure 3) which suggests that the wild-type forms a greater number of electrostatic interactions at the DNA interface than the mutant. The thermodynamics of binding also confirms this result. While the enthalpy of binding is more thermodynamically favourable for the mutant, the electrostatic entropy component is more favourable for the wild-type. This implies that although the A539P mutant has a larger binding interface and forms more sequence specific contacts with the DNA

bases, the wild-type forms more ionic contacts with the DNA phosphate backbone, which are likely the cause of the higher DNA affinity reported for the wild-type (Morris and Fanucchi, 2016).

It is clear that both specific interactions with the DNA bases and non-specific interactions with the phosphate backbone occur upon DNA binding. Although we show here that the A539P mutant may have a higher specificity of binding than the wild-type (due to the decreased number of ionic contacts and greater heat capacity of binding), both wild-type and mutant do have some degree of specificity for their targets as they do not interact with a random sequence (Webb et al., 2017). Specific site recognition and formation of a stable protein-DNA complex relies on two requirements. One is hydrogen bonds formed between the residues of the protein recognition motif and the nitrogenous bases which contributes to the enthalpic binding term and change in heat capacity (direct readout). The other is shape complementarity and conformational adjustments to maximise electrostatic interactions between the DNA backbone and the surrounding protein structure through shape and charge complementarity (indirect readout) (Steffen et al., 2002). In agreement with the heat capacity studies (Figure 2), the entropic component associated with conformational changes is the minor contributing factor to the entropy of binding (Figure 3). Therefore in order to form site specific interactions with the DNA, both the wild-type and the A539P FOXP2 FHD depend on electrostatic and hydrogen bonds with the backbone of the DNA target sequence and flanking regions (direct readout) rather than large conformational changes. The nuclear magnetic resonance (NMR) structures of the FHD of FOXP1 and FOXO4 agree with this observation, having a stable folded structure in solution in the absence of DNA (Weigelt et al., 2001; Chu et al., 2011).

Chimeric protein and NMR studies have shown that the region spanning helix-2 and the hinge loop is known to be a significant contributor to the specificity of the FHD (Overdier et al., 1994; Pierrou et al., 1994; Marsden et al., 1998). The data obtained in this study corroborates those findings by showing that the protein-DNA interactions that occur between the FHD and the consensus binding site differ significantly due to a point mutation in the hinge loop region. The difference in the number and types of interactions has the potential to alter the specificity of the FHD by altering the depth of groove insertion (and so the number of base specific hydrogen bonds) as well as facilitating novel backbone ionic contacts that would otherwise be restricted by a rigid hinge loop.

It is interesting, considering the highly conserved nature of the recognition helix, that the various members of the FOX family are able to recognise different specific sequences. According to our results, it is possible that specificity is linked to backbone flexibility. One of the regions that appears to have a remarkable change in backbone flexibility upon binding of the wild-type to DNA is the wing region, located directly on the C-terminal to the recognition helix (Figure 4). Of all the regions in the FOX FHD, the β-sheet wing is one of the least conserved. Interestingly, the majority of known FHD structures reveal that there are contacts made between the wing and the DNA largely between positively charged residues and the phosphate backbone (Stroud et al., 2006; Tsai et al., 2007; Littler et al., 2010). Considering both the flexibility and the variability of this region, these contacts may be key to distinguishing binding site selectivity between FOX family members. Perhaps the wing acts to regulate the activity of the FOX proteins through protein-protein interactions with other transcription factors as observed in the RHR domain of NFAT and the FOXP2 FHD co-crystal structure (Wu et al., 2006). This could provide an additional mechanism to alter gene regulation by the various FOX transcription factors despite the highly conserved recognition sequence of the FOX family. Recently, several studies have aimed to disrupt the DNA binding of FOXM1 with specific small molecule compounds due to its role in the progression and survival of several cancers (Gormally et al., 2014; Marsico and Gormally, 2015). The flexibility and variability of the wing region among FOX family members may prove an ideal target for further drug studies, both to provide specificity towards FOX proteins and possibly even to control the coordination of specific gene expression regimes by disrupting necessary protein-protein interactions (Fontaine et al., 2015).

The role of the FHD hinge loop *in vivo*

Regulation of gene expression depends on the cellular context and many transcription factors can act to either increase or reduce appropriate gene targets by coalescing multiple cellular signals through macromolecular complex formation and post-translation modification. Indeed, FOXP2 acts as both a transcriptional repressor and activator depending on the gene target (Spiteri et al., 2007). The reporter assay described here shows the transcriptional activation capability of FOXP2 for a cognate promoter sequence. The study was performed on full length FOXP2 wild-type and full length FOXP2 A539P mutant. Although the A539P mutation has been shown to abolish dimerisation of the isolated FHD *in vitro* (Stroud et al., 2006; Chu et al., 2011), it cannot be said for certain

whether the mutation alters the quaternary structure of full length FOXP2 *in vivo*. It is very likely that full length A539P FOXP2 is still capable of forming dimers if only via the leucine zipper interface. It is probable, though, given the evidence from the isolated FHD, that the A539P mutation will prevent dimerisation of the full length protein at the FHD dimer interface *in vivo*.

The results from the study show that FOXP2 wild-type is more transcriptionally active than the A539P mutant; however, the hinge-loop mutation does not completely inactivate FOXP2 activity (Figure 6). Therefore, the ability of the FHD to form dimers may not be absolutely essential for transcriptional activity of FOXP2 although it may be required for optimal activity. This hints at the idea of dimerisation of the FOXP2 FHD acting as a regulatory mechanism. From the results of this study, we have shown that wild-type has a more flexible backbone than the mutant and it has the ability to associate with the DNA with a more dynamic open structure. The fact that wild-type is more transcriptionally active than the mutant implies that backbone flexibility and loose association with the DNA may be an important consequence of the evolutionary hinge loop mutation unique to the FOXP subfamily members by improving DNA affinity (increasing occupancy time at the promoter) as well as by providing a point for domain-swapping to occur at the appropriate promoter site. It is certainly possible that the wild-type assumes the monomeric form when locating and binding the consensus site in a more open conformation after which it can form hetero- or homo-typic domain-swapped dimers with other FOXP subfamily member FHDs in a context and sequence dependent manner. Precisely what effect FHD dimerisation could have on multi-component transcriptional complex formation, aggregation of proximal promoter sites or how it is controlled remains unclear. The quaternary state of the FOXP2 FHD may alter the recruitment of transcription factors necessary for appropriate target gene regulation. For example, NFAT has been shown to interact directly with the monomeric FOXP2 FHD (Wu et al., 2006). The role of FOXP2 FHD dimerisation in protein-protein interactions with NFAT and other transcription factors is an important avenue for future studies.

Conclusions

The members of the FOX family of transcription factors have become clinically relevant targets in recent years owing to their overarching involvement in the progression and survival of a cancerous cell state in several tissues. Understanding the structural biology of specific protein-DNA interactions is essential to the design of novel therapeutics that target and control the transcriptional master regulators of a diseased cell state. Here, we present data that shows a distinct difference in the DNA binding capacity between the wild-type and the A539P mutant, FOXP2 FHD. We demonstrate that the inherent change in dynamics of the hinge-loop region (the region spanning helix-2 and helix-3) in the FOXP2 FHD regulates DNA binding both *in vitro* and *in vivo* by altering the global dynamics of the FOXP2 FHD, highlighting the critical role of the hinge loop in protein dynamics and in altering DNA binding characteristics of closely related transcription factor family members. We show that the wild-type has a more flexible backbone structure which allows it to make more electrostatic contacts with the DNA than the A539P mutant, however, the majority of these interactions are non-specific and furthermore it inserts into the DNA more shallowly than the mutant. This loose binding appears to be necessary for efficient transcription. Overall these results indicate that the hinge-loop connecting helix-2 and helix-3 plays a significant role in regulating the mechanism of DNA binding of the FHD. This is an important distinguishing factor between FOXPs and other FOX proteins that do not have the flexible hinge-loop.

Materials and methods

Protein expression and purification

T7 *Escherichia coli* bacteria were transformed with a pET-11a plasmid housing the coding sequence for human FOXP2 FHD (residues 501–584) fused with an N-terminal hexahistidine tag under the control of an IPTG inducible promoter. Site-directed mutagenesis was performed to generate the A539P mutation using the Quickchange site-directed mutagenesis kit (Agilent, Santa Clara, CA, USA). Cultures were grown to an OD_{600} of 0.6–0.8 at 37°C in an aerobic shaker rotating at ~200 rpm. Once the appropriate optical density was reached the cultures were cold shocked at 4°C for 10 min before 0.5 mM IPTG was added to induce heterologous protein expression. The culture was then incubated at 20°C for 20 h in an aerobic shaking incubator at 200 rpm to allow for sufficient soluble heterologous protein expression.

Immobilised metal ion affinity chromatography (IMAC) was utilised for coarse purification of the FOXP2 FHD-fusion protein from the bacterial cell milieu (Supplementary Figure 1). The IMAC column was equilibrated with 10 column volumes of equilibration buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 25 mM imidazole). The cell lysate was then loaded onto the column at a flow rate of 2 ml/min. The column was washed with 10 column volumes of high salt buffer (20 mM Tris-HCl pH 7.5, 1.5 M NaCl, 25 mM imidazole) before the bound protein was eluted in a single step with elution buffer (20 mM Tris-HCl, 500 mM NaCl, 300 mM imidazole). The N-terminal His-tag was removed by incubation of the protein with thrombin for 4 h at 20°C. The protein was further purified

using successive rounds of IMAC, benzamidine affinity chromatography and size-exclusion chromatography as described elsewhere (Blane and Fanucchi, 2015; Morris and Fanucchi, 2016).

Oligonucleotides

The following duplex cognate DNA sequence TTAGG**TGTTTAC**TTTCATAG containing a single binding site (in bold) has been shown to have a strong affinity for the FOXP2 FHD (Nelson et al., 2013; Morris and Fanucchi, 2016; Webb et al., 2017) and was used exclusively in this study. Duplex DNA was synthesised at Integrated DNA Technology, South Africa and was prepared to a stock concentration of 200 μM in TBE buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA, pH ~ 8.3). The determination of DNA concentration was performed in triplicate using ultraviolet (UV)-absorbance and the average was taken as the final concentration for downstream experiments.

Heat capacity of DNA binding

The heat capacity of binding for the wild-type and A539P FOXP2 FHD was determined using isothermal titration calorimetry on a NanoITC instrument (TA Instruments, NJ, USA; Supplementary Table 1). The heat capacity value was obtained from the gradient of a linear regression fitted to the enthalpies of a series of five titrations performed at five temperatures in the range of 10–30°C. A typical titration consisted of 5 μl injections of 70–100 μM protein into 7–10 μM DNA. Both the protein and DNA were dialysed thoroughly against the same binding buffer (10 mM HEPES pH 7.5, 100 mM NaCl and 0.02% NaN_3). Blank titrations of protein into buffer and buffer into DNA were performed to account for heats of dilution. Enthalpies were determined from a nonlinear independent sites regression using the NITPIC software package (Keller et al., 2012). Each titration experiment was performed at least in duplicate and the replicates were averaged. Errors are the standard deviation of the averaged replicates.

Salt effects on DNA binding

The electrostatic contributions to DNA binding of both the wild-type and the A539P mutant were calculated by performing DNA binding experiments at increasing salt concentrations and applying the following equation (Privalov et al., 2011; Supplementary Table 1):

$$\log(K_a) = -N \cdot \log[\text{Salt}] + \log(K_{\text{nel}})$$

The number of electrostatic contacts formed between the protein and the DNA (N) and the non-electrostatic association constant (K_{nel}) were determined by fitting a linear regression to the double log plot of salt concentration and the corresponding experimental K_a. The association constants (K_a) for DNA binding were obtained from non-linear regression of an independent sites model fitted to isotherms determined by isothermal titration calorimetry at increasing salt (NaCl) concentrations. K_{nel} was calculated at a 1 M salt concentration. The entropy was dissected into contributions from the conformational changes and counterion exclusion using the Gibbs-Helmholtz equation. The entropy at 293 K was chosen as the standard to compare the wild-type FOXP2 FHD to the A539P FOXP2 FHD. The number of ionic contacts with the DNA backbone were estimated from N by considering the DNA phosphate counter-ion occupancy of NaCl, $\psi = 0.64$ (Olmsted et al., 1995; Privalov et al., 2011).

Isothermal (293 K) titrations of 70–100 μM wild-type or 70–100 μM A539P FOXP2 FHD into 7–10 μM DNA were performed at 293.15 K in binding buffer (10 mM HEPES pH 7.5 and 0.02% NaN_3) adjusted to one of five salt concentrations in the range of 50–150 mM NaCl. Each titration consisted of 40 5 μl injections of protein into 950 μl cognate DNA using a NanoITC instrument (TA Instruments, NJ, USA). Each titration was performed at least in duplicate and the replicate values were averaged. Errors are the standard deviation of the averaged values. Each binding isotherm was fit with a non-linear independent sites model minimising the chi-squared between the model and experimental data points.

Dynamic quenching studies

The Stern-Volmer coefficient (K_{SV}) describes the solvent accessibility of the tryptophan residues based on the degree to which they are quenched by increasing concentrations of quencher (Q) in solution (Eftink and Ghiron, 1976):

$$\frac{F_0}{F} = 1 + K_{SV} \cdot [Q]$$

The Stern-Volmer coefficient (K_{SV}) was determined for the apo and DNA bound form of the wild-type and A539P FOXP2 FHD (Supplementary Figure 9). Fluorescence measurements were performed on a Jasco FP-6300 fluorimeter with an excitation wavelength of 295 nm. Samples consisted of 2 μM FOXP2 FHD in binding buffer (10 mM HEPES pH 7.5, 100 mM NaCl and 0.02% NaN_3) with increasing acrylamide (quencher) concentrations from 0 to 250 mM. For DNA bound studies 2 μM FOXP2 FHD was incubated with an equivalent of cognate DNA oligonucleotide for 30 min at 20°C prior to performing the experiments. Buffer only and DNA-buffer blanks were subtracted before data analysis. All experiments were performed in triplicate and averaged following analysis. Errors are the standard deviation of the averaged data.

Hydrogen-deuterium exchange mass spectrometry

The *in vitro* structural dynamics of the apo and DNA bound forms of wild-type and A539P FOXP2 FHD were studied by hydrogen-deuterium exchange mass spectrometry (Supplementary Figures 7 and 8). Labelling, quenching and proteolytic cleavage experiments were performed on a PAL HDX system (LEAP Technologies, USA) coupled to an Agilent 1100 high-poser liquid chromatography (HPLC) system (Agilent, USA). Mass spectrometry was performed on an AB Sciex 6600 TripleTOF mass spectrometer (AB Sciex, USA). Protein labelling consisted of a 10–3600 s incubation of 20–30 μg of wild-type or A539P FOXP2 FHD, with and without an equivalent mass of cognate DNA oligonucleotide, in binding buffer (10 mM HEPES pH 7.5, 100 mM NaCl) made up with 90% D_2O at 293 K. Samples were then quenched by a two-fold dilution in quench buffer (20 mM phosphate pH 4.5, 100 mM TCEP and 3.4 M guanadine-HCl) held at 0°C. The protein was then fragmented by incubation on an inline Porozyme immobilised pepsin chromatography column

(Life Technologies) at 4°C for 30 s before being desalted on an Acclaim PepMap trap column (0.3 × 5 mm) and subsequently loaded onto a Kinetex C_{18} reverse phase chromatography column (Phenomenex, USA). Peptides were separated onto the mass spectrometer at a flow rate of 200 µl/min with an elution gradient of 5–40% buffer B (80% ACN/0.1% FA).

Peptide mass analysis was performed on an AB Sciex 6600 TripleTOF in Data Dependent Acquisition (DDA) mode. In DDA mode, precursor scans were acquired from *m/z* 360–1500 using an accumulation time of 250 ms followed by 30 product scans, acquired from *m/z* 100–1800 at 100 ms each, for a total scan time of 3.3 s. Charge ions, falling between 1^+ and 5^+, were automatically fragmented in the Q2 collision cell using nitrogen gas. The collision energy was chosen automatically based on the *m/z* and the charge. Peptide identification was performed in PEAKS 6 (Bioinformatics Solutions Inc., USA). The degree of deuteron incorporation was determined with the HDXaminer software package (Sierra Analytics, USA).

Luciferase reporter assays

HEK293 cell cultures were plated at a density of 1×10^4 cells per well in a 96-well plate and grown to confluency in antibiotic free Dulbecco's Modified Eagle's Medium (DMEM) medium at 37°C with 5% CO_2. Each well was transfected with 1.6 µg of pcDNA4 vector containing the full length wild-type FOXP2 coding sequence or the full length A539P FOXP2 coding sequence obtained by site-directed mutagenesis of the wild-type coding sequence using the Quikchange Lightening site-directed mutagenesis kit followed as per manufacturer's instructions (Agilent, USA; Supplementary Figure 11). Negative controls were performed by the addition of a transfection reagent, transfection with 1.6 µg pGL4 vector under the control of a 6X FOXP2 consensus sequence synthesised by Integrated DNA Technology (Cape Town, South Africa; Supplementary Figure 10) and transfection with 1.6 µg pRL-TK vector encoding *Renilla* firefly luciferase under control of a cognate tyrosine kinase promoter. Transfections were performed using Fugene HD transfection reagent as per the manufacturer's instructions (Promega). Transfected cells were incubated for a further 24 h. Luciferase assays were performed using the Dual-Glo luciferase assay kit followed as per the manufacturer's instructions (Promega). Luminescence readings were taken on a GLoMax 96 Microplate Luminometer (Promega). Transfection efficiency was normalised by co-transfection with a pRL-TK vector encoding *Renilla* luciferase under control of a tyrosine kinase promoter and by taking the ratio of the firefly luciferase activity to *Renilla* luciferase activity. Replicates of at least three were performed for each FOXP2 sequence and control.

Acknowledgements: We would like to acknowledge Kerry Hanmer and Dr. Demetra Mavri-Damelin from the School of Molecular and Cell Biology at the University of the Witwatersrand for their assistance in culturing the HEK293 cells used in the dual luciferase reporter assays. We also extend our gratitude to Dr. Sonja Vernes (Max Planck Institute for Psycholinguistics, The Netherlands) for the generous gift of the pcDNA4-FOXP2 vector used in the dual luciferase reporter assays.

Funding: This work was supported by the University of the Witwatersrand; South African National Research Foundation, Funder Id: 10.13039/501100001321 (Grant 80681 to S.F., 68898 to H.W.D.); the South African Research Chairs Initiative of the Department of Science and Technology (Grant 64788 to H.W.D.) and the Medical Research Council of South Africa, Funder Id: 10.13039/501100001342.

References

Bandukwala, H.S., Wu, Y., Feuerer, M., Chen, Y., Barboza, B., Ghosh, S., Stroud, J.C., Benoist, C., Mathis, D., Rao, A., et al. (2011). Structure of a domain-swapped FOXP3 dimer on DNA and its function in regulatory T cells. Immunity *34*, 479–491.

Blane, A. and Fanucchi, S. (2015). Effect of pH on the structure and DNA binding of the FOXP2 forkhead domain. Biochemistry *54*, 4001–4007.

Chu, Y.P., Chang, C.H., Shiu, J.H., Chang, Y.T., Chen, C.Y., and Chuang, W.J. (2011). Solution structure and backbone dynamics of the DNA-binding domain of FOXP1: insight into its domain swapping and DNA binding. Protein Sci. *20*, 908–924.

Clark, K.L., Halay, E.D., Lai, E., and Burley, S.K. (1993). Co-crystal structure of the HNF-3/forkhead DNA-recognition motif resembles histone H5. Nature *36*, 412–420.

Cuiffo, B.G., Campagne, A., Bell, G.W., Lembo, A., Orso, F., Lien, E.C., Bhasin, M.K., Raimo, M., Hanson, S.E., Marusyk, A., et al. (2014). MSC-regulated microRNAs converge on the transcription factor FOXP2 and promote breast cancer metastasis. Cell Stem Cell *15*, 762–774.

deHaseth, P.L., Gross, C.A., Burgess, R.R., and Record Jr., M.T. (1977). Measurement of binding constants for protein-DNA interactions by DNA-cellulose chromatography. Biochemistry *16*, 4777–4783.

Eftink, M.R. and Ghiron, C.A. (1976). Exposure of tryptophanyl residues in proteins. Quantitative determination by fluorescence quenching studies. Biochemistry *15*, 672–680.

Fontaine, F., Overman, J., and François, M. (2015). Pharmacological manipulation of transcription factor protein-protein interactions: opportunities and obstacles. Cell Regen. (London) *4*, 2.

Fujita, E., Tanabe, Y., Shiota, A., Ueda, M., Suwa, K., Momoi, M.Y., and Momoi, T. (2008). Ultrasonic vocalization impairment of Foxp2 (R552H) knockin mice related to speech-language disorder and abnormality of Purkinje cells. Proc. Natl. Acad. Sci. USA *105*, 3117–3122.

Gajiwala, K.S. and Burley, S.K. (2000). Winged helix proteins. Curr. Opin. Struct. Biol. *10*, 110–116.

Gong, X., Jia, M., Ruan, Y., Shuang, M., Liu, J., Wu, S., Guo, Y., Yang, J., Ling, Y., Yang, X., et al. (2004). Association between the FOXP2 gene and autistic disorder in the Chinese population. Am. J. Med. Genet. B Neuropsych. *127B*, 113–116.

Gormally, M.V., Dexheimer, T.S., Marsico, G., Sanders, D.A., Lowe, C., Matak-Vinkoviä, D., Michael, S., Jadhav, A., Rai, G., Maloney, D.J., et al. (2014). Suppression of the FOXM1 transcriptional programme via novel small molecule inhibition. Nat. Commun. *5*, 5165.

Hachigian, L.J., Carmona, V., Fenster, R.J., Kulicke, R., Heilbut, A., Sittler, A., Pereira de Almeida, L., Mesirov, J.P., Gao, F., Kolaczyk, E.D., et al. (2017). Control of Huntington's disease-associated phenotypes by the striatum enriched transcription factor Foxp2. Cell Rep. *21*, 2688.

Hannenhalli, S. and Kaestner, K.H. (2009). The evolution of Fox genes and their role in development and disease. Nat. Rev. Genet. *10*, 233–240.

Kaestner, K.H., Knöchel, W., and Martínez, D.E. (2000). Unified nomenclature for the winged helix/forkhead transcription factors. Genes Dev. *14*, 142–146.

Keller, S., Vargas, C., Zhao, H., Piszczek, G., Brautigam, C.A., and Schuck, P. (2012). High-precision isothermal titration calorimetry with automated peak-shape analysis. Anal. Chem. *84*, 5066–5073.

Lai, E., Clark, K.L., Burley, S.K., and Darnell, J.E. (1993). Hepatocyte nuclear factor 3/fork head or "winged helix" proteins: a family of transcription factors of diverse biologic function. Proc. Natl. Acad. Sci. USA *90*, 10421–10423.

Lai, C.S.L., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F., and Monaco, A.P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. Nature *413*, 519–523.

Li, S., Weidenfeld, J., and Morrisey, E.E. (2004). Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. Mol. Cell Biol. *24*, 809–822.

Littler, D.R., Alvarez-Fernández, M., Stein, A., Hibbert, R.G., Heidebrecht, T., Aloy, P., Medema, R.H., and Perrakis, A. (2010). Structure of the FoxM1 DNA-recognition domain bound to a promoter sequence. Nucleic Acids Res. *38*, 4527–4538.

Liu, C.-C., Richard, A.J., Datta, K., and LiCata, V.J. (2008). Prevalence of temperature-dependent heat capacity changes in protein-DNA interactions. Biophys. J. *94*, 3258–3265.

Manning, G.S. (1978). The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. Q. Rev. Biophys. *11*, 179.

Marsden, I., Jin, C., and Liao, X. (1998). Structural changes in the region directly adjacent to the DNA-binding helix highlight a possible mechanism to explain the observed changes in the sequence-specific binding of winged helix proteins. J. Mol. Biol. *278*, 293–299.

Marsico, G. and Gormally, M.V. (2015). Small molecule inhibition of FOXM1: how to bring a novel compound into genomic context. Genomics Data *3*, 19–23.

Morris, G. and Fanucchi, S. (2016). A key evolutionary mutation enhances DNA binding of the FOXP2 forkhead domain. Biochemistry *55*, 1959–1967.

Murphy, K.P. and Freire, E. (1992). Thermodynamics of structural stability and cooperative folding behavious in proteins. Adv. Protein Chem. *43*, 313–361.

Nelson, C.S., Fuller, C.K., Fordyce, P.M., Greninger, A.L., Li, H., and DeRisi, J.L. (2013). Microfluidic affinity and ChIP-seq analyses converge on a conserved FOXP2-binding motif in chimp and human, which enables the detection of evolutionarily novel targets. Nucleic Acids Res. *41*, 5991–6004.

Olmsted, M.C., Bond, J.P., Anderson, C.F., and Record, M.T. (1995). Grand canonical Monte Carlo molecular and thermodynamic predictions of ion effects on binding of an oligocation (L8+) to the center of DNA oligomers. Biophys. J. *68*, 634–647.

Overdier, D.G., Porcella, A., and Costa, R.H. (1994). The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino-acid residues adjacent to the recognition helix. Mol. Cell Biol. *14*, 2755–2766.

Pierrou, S., Hellqvist, M., Samuelsson, L., Enerbäck, S., and Carlsson, P. (1994). Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending. EMBO J. *13*, 5002–5012.

Prabhu, N.V. and Sharp, K.A. (2005). Heat capacity in proteins. Annu. Rev. Phys. Chem. *56*, 521–548.

Privalov, P.L., Dragan, A.I., and Crane-Robinson, C. (2011). Interpreting protein/DNA interactions: distinguishing specific from non-specific and electrostatic from non-electrostatic components. Nucleic Acids Res. *39*, 2483–2491.

Record, J.M.T., Anderson, C.F., and Lohman, T.M. (1978). Thermodynamic analysis of ion effects on binding and conformational equilibria of proteins and nucleic-acids – roles of ion association or release, screening, and ion effects on water activity. Q. Rev. Biophys. *11*, 103–178.

Rousseau, F., Schymkowitz, J.W., and Itzhaki, L.S. (2003). The unfolding story of three-dimensional domain swapping. Structure *11*, 243–251.

Sin, C., Li, H., and Crawford, D.A. (2014). Transcriptional regulation by FOXP1, FOXP2, and FOXP4 dimerization. J. Mol. Neurosci. *55*, 437–448.

Spaniel, F., Horacek, J., Tintera, J., Ibrahim, I., Novak, T., Cermak, J., Klirova, M., and Hoschl, C. (2011). Genetic variation in FOXP2 alters grey matter concentrations in schizophrenia patients. Neurosci. Lett. *493*, 131–135.

Spiteri, E., Konopka, G., Coppola, G., Bomar, J., Oldham, M., Ou, J., Vernes, S.C., Fisher, S.E., Ren, B., and Geschwind, D.H. (2007). Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. Am. J. Hum. Genet. *81*, 1144–1157.

Spolar, R. and Record, M. (1994). Coupling of local folding to site-specific binding of proteins to DNA. Science *263*, 777–784.

Steffen, N.R., Murphy, S.D., Tolleri, G.W., and Lathrop, R.H. (2002). DNA sequence and structure: direct and indirect recognition in protein-DNA binding. Bioinformatics *18*, S22–S30.

Stroud, J.C., Wu, Y., Bates, D.L., Han, A., Nowick, K., Paabo, S., Tong, H., and Chen, L. (2006). Structure of the forkhead domain of FOXP2 bound to DNA. Structure *14*, 159–166.

Takahashi, H., Takahashi, K., and Liu, F.-C. (2009). FOXP genes, neural development, speech and language disorders. In: Advances in Experimental Medicine and Biology, K. Maiese, ed. (Berlin: Springer), pp. 117–129.

Teramitsu, I. (2004). Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. J. Neurosci. *24*, 3152–3163.

Tsai, K.L., Sun, Y.J., Huang, C.Y., Yang, J.Y., Hung, M.C., and Hsiao, C.D. (2007). Crystal structure of the human FOXO3a-DBD/DNA complex suggests the effects of post-translational modification. Nucleic Acids Res. *35*, 6984–6994.

Webb, H., Steeb, O., Blane, A., Rotherham, L., Aron, S., Machanick, P., Dirr, H., and Fanucchi, S. (2017). The FOXP2 forkhead domain binds to a variety of DNA sequences with different rates and affinities. J. Biochem. *162*, 1–10.

Weigelt, J., Climent, I., Dahlman-Wright, K., and Wikstrom, M. (2001). Solution structure of the DNA binding domain of the human forkhead transcription factor AFX (FOXO4). Biochemistry *40*, 5861–5869.

Wong, K.K., Gascoyne, D.M., Soilleux, E.J., Lyne, L., Spearman, H., Roncador, G., Pedersen, L.M., Møller, M.B., Green, T.M., and Banham, A.H. (2016). FOXP2-positive diffuse large B-cell lymphomas exhibit a poor response to R-CHOP therapy and distinct biological signatures. Oncotarget *7*, 52940–52956.

Wu, Y., Borde, M., Heissmeyer, V., Feuerer, M., Lapan, A.D., Stroud, J.C., Bates, D.L., Guo, L., Han, A., Ziegler, S.F., et al. (2006). FOXP3 controls regulatory T cell function through cooperation with NFAT. Cell *126*, 375–387.

Supplemental Material: The online version of this article offers supplementary material (https://doi.org/10.1515/hsz-2018-0185).