Citation for published version (APA):
T-bet as a key regulator of mucosal immunity.

Rami Mohamed\textsuperscript{1, 2}, Graham M. Lord\textsuperscript{1, 2*}

1 Department of Experimental Immunobiology, Division of Transplantation Immunology and Mucosal Biology, King’s College London, London, United Kingdom.

2. National Institute for Health Research Biomedical Research Centre, Guy's and St Thomas' National Health Service Foundation Trust, London, United Kingdom.

*Corresponding Author:

Professor Graham Lord

LORD Lab

Department of Nephrology and Transplantation

5th Floor Tower Wing,

Guy's Hospital,

Great Maze Pond,

London,

SE1 9RT.

United Kingdom.

Email: graham.lord@kcl.ac.uk

Key Words: T-bet, TBX21, mucosal immunity, mucosal homeostasis
Abbreviations

CCR1  C-C chemokine receptor type 1
ChIP-seq  Chromatin Immunoprecipitation with Sequencing
cNK  Conventional NK cells
GATA-3  GATA-binding protein 3
ID1  Inhibitor Of DNA Binding 1
IFNγ  Interferon gamma
IL  Interleukin
ILC  Innate Lymphoid Cell
LIN-  Lineage Negative (CD4- CD3- NK1.1- GR-1- CD11b-)
NK cell  Natural Killer Cell
RAG  Recombination Activation Gene
RORγT  Retinoic acid orphan receptor gamma
RUNX1  Runt-related transcription factor 1
T-BET  T-box factor expressed in T-cells
TBR1  T-box, brain, 1
TBX21  T-box 21
Th9  T Helper 9
ThPOK  Th inducing POZ-Kruppel Factor
T-reg  Regulatory T-cell
IEL  Intraepithelial Lymphocyte
ILC  Innate Lymphoid Cells
TNFR  Tumour Necrosis Factor Receptor
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNF</td>
<td>Tumour Necrosis Factor</td>
</tr>
<tr>
<td>TRUC</td>
<td>T-bet&lt;sup&gt;−/−&lt;/sup&gt; RAG&lt;sup&gt;−/−&lt;/sup&gt; Ulcerative Colitis</td>
</tr>
<tr>
<td>CpG</td>
<td>Cytosine-phosphate-Guanine</td>
</tr>
<tr>
<td>EAE</td>
<td>Experimental Autoimmune Encephalomyelitis</td>
</tr>
<tr>
<td>EOMES</td>
<td>Eomesodermin</td>
</tr>
</tbody>
</table>
Abstract

Initially understood to be a key regulator of IFNγ producing helper T-cells, our knowledge of T-bet’s functional roles has expanded considerably to encompass a growing range of cellular lineages. In addition to regulating other IFNγ producing adaptive immune cells, it is now apparent that T-bet plays a fundamental role in the regulation of innate immune responses across mucosal surfaces. This homeostatic role is highlighted by the spontaneous colitis that occurs when T-bet is deleted from innate immune cells in RAG−/− mice. Using this model as a focal point, we review our understanding of T-bet’s regulation of adaptive and innate immune systems, focusing particularly on mucosal populations including innate lymphoid cells, dendritic cells, and intraepithelial lymphocytes. With the increasingly diverse effects of T-bet on different lineages, the classical binding-centric paradigm of T-bet’s molecular functionality has increasingly struggled to account for the versatility of T-bet’s biological effects. Recent recognition of the synergistic interactions between T-bet and other canonical transcription factors has led to a co-operative paradigm that has provided greater explanatory power. Synthesising insights from ChIP-seq studies and evolutionary biology, we expand the co-operative paradigm further and suggest a network approach as a powerful way to understand and model T-bet’s diverse functionality.
Introduction

The T-box family of proteins are an ancient, evolutionary conserved, family of transcription factors with key roles in development mediated through highly specific patterns of expression that are tightly regulated through both time and space (1,2). T-BOX21 (TBX21), or T-box factor expressed in T-cells (T-bet), was first cloned, in 2000, as a transcription factor specific to interferon-gamma (IFNγ) producing helper T-cells that could account for the subtype specific production of this cytokine (3). Since this original description, an exponential growth in T-bet related research has solidified our understanding of T-bet as a key regulator of adaptive immunity.

In recent years, however, it is increasingly clear that T-bet has wide ranging functions outside of helper T-cell biology (4). In particular, in the setting of the mucosal surfaces, T-bet has been found to regulate the developmental pathways of a wide variety of both innate and adaptive immune cells, including innate lymphoid cells (ILC) and intraepithelial lymphocytes (IEL). Alongside this emerging role in the ontogeny of several cellular lineages, T-bet has also been found to actively regulate innate mucosal immune responses as demonstrated by the spontaneous communicable ulcerative colitis which develops in the setting of T-bet deficiency, in mice lacking adaptive immune cells (5).

As our knowledge of T-bet’s diverse cell-specific functions has grown, and with the discovery of innate lymphoid cells, a paradigm has emerged where T-bet is now conceptualised not only as a ‘master regulator’ of T helper 1 (Th1) cells but as a broad, conserved regulator, of type 1 immune responses (4). A picture that is supported by the general conservation of type 2 and type 3 responses, and their respective transcription factors, GATA-binding protein 3 (GATA-3) and retinoic acid-related orphan receptor
gamma (RORγt), across mammalian evolution (4). Notwithstanding this shift in conceptualisation, our classical understanding of T-bet’s molecular functionality, which we refer to as a binding-centric paradigm, has increasingly struggled to account for the increasingly divergent effects of T-bet in a growing number of immune cells. If T-bet’s binding to select gene promoters and enhancers containing its T-box recognition motif, as classically understand, is ipso facto sufficient to drive specific patterns of gene expression, how is it that T-bet binding can have such contrasting effects in different cellular lineages? To explain this paradox, we synthesise emerging empirical evidence into a network theory of T-bet functional effects, where it is argued that the functionality of this transcription factor is dependent on the network of transcription factors already present within a cell thus explaining how T-bet’s expression can regulate diverse, even opposing, functional programs in different cellular milieus.

There have been many excellent reviews of the role of T-bet in adaptive immune cells (6,7). This review focuses on T-bet’s role in governing mucosal immune responses given the increasing realisation of the unique properties of mucosal immune responses and their relevance to human health. A key focus of the review is on the molecular functions of T-bet that are able to account for the variety of its functional effects within different settings. As most of our understanding of the molecular functionality of T-bet has arisen from studies in helper T-cells, we begin by reviewing what has been learnt from T-bet’s role in helper lymphocytes. Subsequently, we present the T-bet deficient model of colitis in recombination activating gene (RAG) knock-out mice as a powerful model for capturing the interconnected functions of T-bet on the innate immune landscape of the intestinal mucosa. We explore the role of T-bet in key mucosal populations such as innate lymphoid cells, intraepithelial lymphocytes and dendritic cells. These cellular
lineages are highlighted, in particular, as they showcase the contrasting ways in which T-bet can regulate diverse cell fates. These contrasting roles are used as a background to introduce the network theory as an adaptation of the co-operative paradigm increasingly popular in molecular immunology. We discuss the therapeutic potential of targeting T-bet, and finally summarise key outstanding questions in the molecular biology of T-bet.
T-bet as a 'master-regulator' of helper T-cells and the emergence of the classical paradigm of immune transcriptional regulation.

The discovery by Mossman and Coffman of distinct types of helper T-cells, established one of the central paradigms of immunology: that adaptive immune responses in addition to being quantitatively tuned to antigen, through clonal selection, are also qualitatively tuned through the selection of unique effector programs (8). Despite the accumulation of substantial in vivo evidence supportive of this paradigm, it would take 10 years before a molecular explanation could be provided to account for the unique effector programs of helper T-cells. The identification of GATA-3, in 1997, followed by T-bet, three years later, as subtype specific transcription factors established a paradigm of ‘master-regulators’ controlling specialised cell fates (3,9). This mode of thinking has deeply influenced immunologists, as illustrated by the reluctance to fully accept newly proposed lineages such as the T helper 22 (Th22) or T helper 9 (Th9) lineages, which lacking a unique transcription factor fail to dovetail smoothly into this paradigm.

Early hypothesis-driven studies into the molecular effects of T-bet focused on key functional loci known to be differentially expressed between helper T-cell subsets, such as IFNγ and interleukins 4 (IL-4) and 5 (IL-5) (10). These early studies, using gene-reporter assays, identified the specific in-cis up-regulation of T helper 1 (Th1) specific cytokines through T-bet binding and the reciprocal inhibition of competing cytokines. A crucial insight, made early in the analysis of T-bet, is that while positive gene regulation arises from T-bet’s interaction with DNA, negative regulation, in many instances, occurs through protein-protein interactions with other canonical immune regulators (11). This negative regulation is illustrated by phosphorylation of GATA-3 by T-bet which
functions to repress the GATA-3 program in Th1 cells partly through the re-direction of GATA-3 from its default binding sites (11).

Thus a model arose whereby the specific binding of T-bet to core promoters was understood to cause up-regulation of a lineage-specific transcriptional program while inhibiting opposing cellular programs, in-part, through protein interactions. So well did the available empirical data fit this model, that it has led historically to a heavily determined view of T-bet, whereby this transcription factor alone was believed to be ‘necessary and sufficient’ for the differentiation of Th1 cells. Central to this paradigm was the view that specific binding patterns of T-bet, as informed by hypothesis driven gene reporter assays, were ipso facto sufficient to account for the transcription factor’s positive genomic effects (6). The development of ChIP-seq, as an unbiased way to assay genome-wide protein binding, appeared to broadly support the main assumptions of this binding-centric paradigm, where genes known to be functionally regulated by T-bet were shown to correlate with T-bet binding to genomic regions possessing the canonical T-box motif (12-14). Additionally, Zhu et al. demonstrated that of 219 genes positively regulated by T-bet in helper T-cells, over 97, including characteristic markers of Th1 identity such as IFNγ, IL12rb and CXCR3, were directly bound by T-bet (15).
Wider roles for T-bet in adaptive immune response and the emergence of T-bet as central regulator of mucosal homeostasis.

With the emergence of the influential Th1/Th2 dichotomy, a paradigm emerged whereby T-bet through transcriptional control of specific cytokines and homing molecules, such as the Th1 specific chemokine (C-X-C motif) receptor 3 (CXCR3), determined cellular fate and functionality (16). The discovery of a role for T-bet in cytotoxic T-cells (CD8+ T-cells) and NK T-cell (NKT) development strengthened this paradigm, as T-bet in these lineages was found to regulate, essentially, the same gene loci as in helper T-cells (11,17). Thus a quantitative expansion in T-bet’s functional roles, reinforced the classical qualitative understanding of how T-bet functions through inducing specific patterns of gene expression through transcriptional control – giving rise to the idea that T-bet is a broadly conserved regulator of type 1, interferon-gamma based immune responses (4).

A significant development in T-bet biology, was the discovery that T-bet could regulate mucosal homeostasis and the homeostatic response to intestinal microbiota (5,18). This, in addition to being a further quantitative expansion, was also a qualitative expansion as T-bet plays a very different role in these cells compared to other adaptive immune cells. The mucosal surfaces of the human body, including those of the respiratory and the digestive tract account for a combined surface area of 300m², and are host to a significant microbial load with over $10^{12}$ bacteria/cm³ in the colon alone (19). The mucosa, therefore, represents the primary site of exposure to non-self, and has a unique immune architecture, or MALT (mucosa associated lymphoid tissue) that facilitates this homeostatic role.
Amongst the earliest evidence for a T-bet specific role in mucosal immune responses is the unexpected development of spontaneous colitis in mice who lack T-bet in the innate immune compartment (20). This T-bet$^{-}$ RAG$^{-}$ Ulcerative Colitis (TRUC) is characterised by inflammation limited to the colon that histologically mimics human ulcerative colitis. Pathology in TRUC arises through a series of well characterised events, summarised in Figure I. The earliest detectable event in TRUC mice is increased Tumour Necrosis Factor (TNF) production by colonic dendritic cells, which is followed by the selective outgrowth of both anaerobic and gram negative organisms in the colonic microbiota. This outgrowth is unique to the TRUC genotype, and does not arise in RAG$^{-}$ mice. Interestingly, mice with the TRUC genotype, but bred in specific-pathogen-free conditions do not develop colitis and are called TRnUC (T-bet$^{-}$ RAG$^{-}$ No Ulcerative Colitis) mice (21). Ribosomal sequencing of TRUC mice, with TRnUCs as controls, has identified an obligatory role for Helicobacter Typhlonius as the microbial species which triggers the cascade of immunopathology (18).

Dendritic cell activation, detectable as early as two weeks results in copious TNF production which is directly attributable to the absence of T-bet (20). Increased TNF produced by CD103$^{+}$ CD11b$^{+}$ dendritic cells has a direct effect on epithelial cells, inducing apoptosis and increased epithelial permeability, though TNF receptor (TNFR) signalling. Autocrine up-regulation of the TNFR on epithelial cells reinforces this pathway. This TNF pathway mediates pathology in the earliest stages of disease which is ameliorated through blockade with an anti-TNF monoclonal antibody. TNF signalling is, however, only pathogenic in the earliest stages of disease and becomes redundant by week 12, as demonstrated by the ineffectiveness of blocking antibodies beyond this time-point (20-22).
Beyond 12 weeks, pathology in TRUC is mediated through, type 3, RORγt+, innate lymphoid cells (ILC3), which produce interleukin 17 (IL-17) and interleukin 22 (IL-22) (23). Both the deletion of ILC3s through anti-CD90 monoclonal antibody, and the abrogation of IL-23 signalling, block disease pathology. CD103\(^+\) CD11b\(^+\) CD11c\(^+\) MHCII\(^+\) dendritic cells are the main source of IL-23 which synergises with TNF to promote the pathogenic release of IL-17 and IL-22 from ILC3s. In addition to the stimulative effect of IL-23 and TNF on ILC3s, the absence of T-bet promotes up-regulation of the interleukin 7 (IL-7) receptor which promotes the survival and longevity of this ILC population. Thus the TRUC model of colitis highlights an essential role for T-bet in regulating mucosal immune responses through cell specific roles in intestinal dendritic cells and innate lymphoid cells.
T-bet’s role in the development and regulation of mucosal immune cells.

A key event in the development of TRUC, is the outgrowth of gram-negative bacteria in the colon. At the time of the original description of TRUC, it was not recognised that T-bet regulates the development of ILC1s and NKp46+ ILC3 populations (24-26). It is thus likely that the selective outgrowth of gram negative bacteria observed in TRUC is a direct consequence of the absence of these T-bet expressing innate populations (20). Recently, it has been observed that T-bet dependent ILC1s, and to a lesser extent NKp46+ ILC3s, play an obligate role in the immunity of RAG-/- mice to* Clostridium difficile,* thus providing supportive evidence in-favour of a cell intrinsic mechanism leading to dysbiosis in TRUC (27). Given the central role of TRUC, in enhancing our understanding of T-bet biology, we briefly review the role of T-bet in the development and function of innate lymphoid cells (ILCs), dendritic cells and intraepithelial cells.

*Innate Lymphoid Cells*

Innate lymphoid cells (ILCs) are CD45+, lineage negative (Lin-) cells that arise from a common lymphoid precursor in the bone marrow (28). Though rare in blood and secondary lymph organs, they are enriched in the mucosal regions of the body including the lungs and intestine and have been found to have important roles in mucosal defence against pathogenic bacteria. ILCs lack antigen-specific receptors, yet share many parallels with helper T-cells (29). The most marked parallel is their subdivision into three groups on the basis of their ability to produce type 1 (IFNγ), type 2 (IL-4, IL-13) and type 3 (IL-17, IL-22) cytokines. In parallel, with helper T-cells: these three groups share a type-specific transcription factor (29). ILCs are of particular excitement, because the
production of specialised cytokines was previously believed to be a unique feature of helper T-cells. ILCs, however, reveal that there is an innate source of these cytokines, which shares conserved mechanisms of transcriptional control - thus leading to a revaluation of some fundamental immunological principles. T-bet, in addition to defining type 1 ILCs, is also expressed in a subset of type 3 innate lymphoid cells (30). The dual expression of T-bet and RORγt in a subset of ILC3s resembles this occurrence in so-called Th17.1 cells which are believed to be the most pathogenic of Th17 cells (31): yet the biological significance of this co-expression in ILCs has yet to be fully determined in vivo.

Type 1 ILCs, initially defined in tonsils and the intestine, are closely related to Natural Killer (NK) cells, through their shared expression of cytotoxicity receptors, shared expression of T-bet and shared ability to produce IFNγ. In contrast to NK cells, they express high levels of the IL-7 receptor (CD127) (32) and are distinguished further from both ILC3s and NK Cells by expression of CD27 (29). A further distinction is that whereas NK cells numbers are unaffected by T-bet deletion, ILC1 development is completely abrogated in T-bet−/− mice: reciprocally, EOMES deletion selectively ablates NK cells with ILC1 development remaining intact (28,30). Thus ILC1s and NK cells are defined by independent developmental-pathways driven by EOMES and T-bet respectively.

The shared possession of cytotoxicity receptors and the dependence on IL-15 signalling by both ILC1s and NK cells, has traditionally obscured their distinct ontogeny: this has been particularly true for a TRAIL+ population of liver NK cells (33). This liver-specific population, in contrast to TRAIL− DX5+ conventional splenic NK cells (cNK)
has a CD49a+ DX5- phenotype, and thus was originally classified as a liver-resident immature NK cell population (34). Careful fate-mapping experiments, however, have shown that like ILC1s, liver-resident NK cells are abrogated with T-bet deletion, and develop independently of EOMES (34,35). Phenotypically, liver-resident NK cells more closely match ILC1s expressing, for example, CXCR6, which is not expressed in cNKs (35). Functionally, liver-resident NK, like ILC1, show less cytotoxic activity in tumour lysis assays in comparison to cNKs (36). Thus, despite heterogenic expression of CD127 and CD27, the preponderance of evidence supports the notion that liver-resident NK cells are best conceptualised as a subset of T-bet dependent ILC1s, distinct from cNKs.

NKp46+ positive ILC3s are RORγt dependent, and arise from NKp46- negative precursors. T-bet is essential for development of this population, and mediates their up-regulation of NKp46. T-bet has also been implicated in the plasticity observed between ILC3 and ILC1 populations (37). In TRUC, the absence of T-bet leads to increased IL-17 and IL-22 production from NKp46- ILC3s which mediates pathology during the later stages of disease. (23). Thus, while there are numerous parallels between helper T-cells and ILCs, in terms of T-bet regulated cytokines, there also subtle differences in terms of gene regulation and subtype specific signatures as suggested by recent transcriptional profiling (38).

Dendritic cells.

Dendritic cells (DC) are key sentinels involved in the early detection of pathogenic insults, and function through modulating the character of subsequent immune activation (39). Dendritic cells are involved in the sampling of antigen, and function
through both the direct polarisation of helper T-cells in situ and also through migration to lymph nodes where they present antigen to T-cells and induce organ specific homing of effector T-cells. As the mucosa represents the primary site of exposure to pathogens, DCs are enriched at these surfaces relative to primary lymphoid organs.

While T-bet is not essential for the development of dendritic cells, a number of functional roles in DC biology have been attributed to T-bet. T-bet deficient splenic DCs (CD11c+ MHCII+) are unable to fully promote T-helper 1, with functional consequences in the setting of *Listeria monocytogenes* infection (40). Furthermore, in immunisation models, T-bet has also been shown to be essential for effective IFNγ responses following immunisation with Toll-Like Receptor (TLR)-9 activating Cytosine-phosphate-Guanine (CpG) nucleotide adjuvants (40). While these roles have been poorly defined at a molecular level, nonetheless, it is clear that within splenic DCs T-bet regulates innate immune and TLR signalling pathways. Furthermore, it is clear that such pathways are implicated in autoimmunity, as T-bet expression in splenic CD11c+ MHC II+ DCs was found to be necessary for rheumatic inflammation in a murine arthritis model, here the absence of T-bet was associated with reduced anti-TNF and Interleukin-1 alpha (IL-1α), as well as perturbed chemokine expression: thus illustrating the effect of T-bet deficiency on the genetic program induced by DC activation.

Equally, the dendritic cell response to immune activation is implicated in TRUC colitis, where T-bet repression of TNF, is unmasked in T-bet- MHCII+ CD11c+ CD11b+ CD103- DCs (22,23). As described above, TNF production from this population initiates immunopathology through a direct effect on epithelial cells and through potentiation of IL-17 production from NKp46- ILC3s. This effect of T-bet deletion on mucosal DCs is
paradoxical as within helper T-cells T-bet is known to activate the TNF locus – thus highlighting the highly specific and unique effects of T-bets within DCs. (10).

**Intraepithelial lymphocytes**

Intraepithelial lymphocytes (IELs) are present alongside epithelial cells, and in the normal human intestine there are as many as 4 IELs for every 10 epithelial cells: as such, IELs represent one of the largest pools of lymphocytes in the human body and alongside intraepithelial CD103+ dendritic cells, they are one of the earliest populations exposed to mucosal insults (41). Though possessing a mature TCR, and thus arguably an adaptive immune cell we have focused on them for their important role in mucosal immunity and for the very unique effects of T-bet within this lineage. As with ILCs, the full biological functions of IELs remain to be fully determined. The pathogenic potential of IELs is emphasised in human coeliac disease, where a numeric up-regulation in IELs is a key pathogenic feature of the disease (42). IELs are also essential mediators of inflammation in the *Toxoplasma gondii* model of intestinal inflammation (43-45). Additionally a novel model of murine intestinal inflammation has been described which is dependent on the transfer of IELs (41).

IELs arise through two routes: central and peripheral. Central IELs are the progeny of committed, triple negative, thymic precursors. Whereas, peripheral IELs arise from the conversion of CD4+ αβ T-cells. The transcription factor, T-bet, has been identified as playing an essential role in the development of both IEL populations, which are reduced in T-bet<sup>−/−</sup> mice (46). Experiments using these mice show that T-bet is essential for the up-regulation of RUNX-3 which then in synergy with T-bet establishes the maturation program of IELs. This maturation program includes the suppression of Th Inducing POZ-Kruppel Factor (ThPOK) which mediates CD4+ cell identity and the up-
regulation of IEL specific genes such as the CD8αα homo-dimer (45). This is significant, as T-bet classically functions in helper T-cells without suppressing ThPOK. Thus, the suppression of this gene by T-bet in IELs represents a specific function that occurs only in this unique setting.
The development of the co-operative paradigm of T-bet functionality and its expansion into a network theory.

The growing biological roles for T-bet necessitate a re-evaluation of our molecular understanding of how T-bet functions. The classical binding-centric paradigm suggests that T-bet positively regulates gene expression through binding to specific genomic sites as recognised by its T-box binding domain. We have called this a binding-centric paradigm, as positive gene regulation is thought to be primarily a function of where T-bet binds in the genome, as is, in turn, determined by the presence of classic T-box motifs in gene promoters and enhancers. This paradigm developed from the empirically determined positive regulation of specific genes such as IFNγ and CD62L in Th1 cells, which were also clearly targeted by T-bet in other adaptive immune cells such as NKT and CD8+ T-cells – suggesting that this positive regulation was a function of T-bet’s innate ability to bind these loci. Yet the idea that the presence of a T-box motif fully determines whether a gene is regulated by T-bet falls apart in view of the recent diametrically opposed roles for T-bet identified in DCs and IELs. We have shown that T-bet upregulates TNF in helper T-cells yet suppresses this same gene in colonic dendritic cells. Furthermore, in IELs, T-bet actively down-regulates ThPOK thus suppressing helper lymphocyte identity, though no such interaction occurs in T-helper one cells. How can the same transcription factor mediate such contrasting effects, given that the T-box binding site in both enhancers and promoters remain consistent between these cells?

Indeed, there are many additional observations that argue against the binding-centric model. These include evolutionary observations and ChIP-seq data. At the evolutionary level it is known that the DNA-binding T-box sequence is highly preserved, with only a few base pair changes in the approximately 180 residues that make up the T-box across great taxonic distances (1,2). Indeed, T-bet, as with other T-BOX family
members, recognises a highly conserved DNA binding motif as has been determined empirically through both competition assays and ChIP-seq. For example, developmental T-BOX proteins such as T-BOX protein in Brain 1 (TBR1) share T-bet’s binding motif, yet regulate very specific developmental pathways. An additional example is Eomesodermin (EOMES), which is also known to recognise a very similar motif to T-bet, yet in the common ID2+ NK precursor (NKP) EOMES drives classical Trial- NK cell development, whereas expression of T-bet drives Trial+ NK cell development, the latter of which homes specifically to the liver and expresses a unique genetic and functional programme (47). Thus the divergent functional effects of T-BOX proteins that recognise highly similar DNA motifs, must result from interactions mediated by the less tightly conserved N and C-terminus of the protein.

ChIP-seq data, in addition, also militates against the binding centric view. While this data has been interpreted to support a binding-centric paradigm, it also allows for a different interpretation. Studies of genomic T-bet binding have consistently shown, that regulated loci account for a small proportion of discernible binding events across a genome. For example, transcriptional profiling availing T-bet+/− mice, has placed the number of genes positively regulated by T-bet at approximately 2 x 10^2 (48). Interestingly, when ChIP-seq is used to analyse T-bet binding sites, the number of T-bet binding sites is over fifty-fold higher, in the region of 1 x 10^4 (10) That is, the vast majority of genomic T-bet binding does not result in transcriptional events. Therefore, as most genomic binding is not accompanied by transcriptional regulation: it is either the character of the promotor, or some other factors present that must determine why some loci are functionally regulated by T-bet while the majority are not effected by T-bet’s presence. Indeed, there is a precedence for such a model in the regulation of T-cell
differentiation, where disinhibition of Inhibitor of DNA-binding protein 1 (ID1) is required for GATA-3 mediated gene expression, despite sufficient binding of GATA-3 to genomic loci – as it is the presence of E-proteins, liberated from sequestration by ID1, at these loci that induces transcription where GATA-3 is present (49).

There has been substantial recognition of the role played by protein-protein interactions between immune lineage determining transcription factors in negative gene regulation. With the expansion of the Th1/Th2 paradigm to encompass Th17 cells, this view has remained robust with T-bet known to negatively regulate Th17 gene expression through interactions with RUNX-1. The increasing recognition of immune plasticity and cells that express multiple lineage transcription factors has given rise to the co-operative paradigm, whereby positive gene regulation is understood to arise from the interaction or co-operation of multiple canonical transcription factors (7). This is powerfully demonstrated in regulatory T-cell (T-reg) biology where the co-expression of T-bet and both GATA-3 alongside FOXP3 has been shown to be essential for optimal suppressive activity in vivo – with T-regs void of these two factors deficient in function (50).

In postulating a network theory we take the co-operative model a step further. We postulate that even in cells expressing a single canonical immune regulator such as T-helper 1 cells or ILC1s, that transcriptional control can best be understood as the output of a network of co-interacting transcription factors, rather than as output of single or small collection of canonical regulators. The attraction of such a theory is that it provides a way to understand how the varied roles of T-bet in-vivo can arise. For example, in dendritic cells, the basal transcription factors within the cell that are distinct from those in helper T-cells are thus likely to explain why in this setting T-bet has unique roles in TLR
signalling and why TNF is uniquely repressed in contrast to other cells where T-bet transactivates this gene.

While it may be tempting to explain away the contradictory roles of T-bet in some cells as the function of hitherto undiscovered protein interactions that are specific to each background lineage, the network theory provides a more systematic and biologically intuitive model that ties together observations from comparative proteomics and ChIP-seq. The strongest evidence for a network theory comes from the most exhaustive study of a canonical immune transcription factor undertaken to-date (51). Wei and colleagues, using ChIP-seq, assayed genome-wide binding of GATA-3 in nine different immune population including the main helper T-cell lineages (Th1, Th2, Th17, natural Treg) and the main thymic T-cell precursors. In line with the T-bet ChIP-seq data presented above, they found that GATA-3 binding induced detectable effects on gene expression on a minute proportion of the total loci bound. Furthermore, as described with T-bet in DCs and IELs, they found that GATA-3 binding at a specific locus could induce either gene activation or repression depending on the cellular context. Additionally they identified a large number of cell specific GATA-3 binding sites, despite conservation of the GATA-3 motif as derived empirically from their data. This substantial cell specificity identified, in even closely related cells, supports our contention that it is the network of factors present with the cells, and not the physical characteristics of the transcription factor, that determines biological outputs.

The network theory is thus a falsifiable model of transcriptional interaction at the far end of a continuum that has evolved from single driver reductionism (master regulator) to one of nuanced nodal interactions (co-operative) to one of sophisticated biological complexity.
Therapeutic targeting of T-bet.

T-bet has been implicated in a large number of inflammatory diseases (52). In the context of IBD, T-bet has been found to be up-regulated in mucosal biopsies of patient’s with Crohn’s disease (53). Given the central role of T-bet in maintaining immune homeostasis, even the attenuation of T-bet activity can lead to disease as emphasised in the TRUC pre-clinical model where the absence of T-bet triggers microbial dysbiosis. Significantly in Asthma, despite no signals from large Genome Wide Association Studies (GWAS) studies, T-bet polymorphisms have been linked repeatedly, through candidate-gene studies, to increased disease severity mediated through an increased disposition to Th2 differentiation (54-56). As such, modulation of T-bet biologically has long been considered an attractive therapeutic target – with both upwards and downwards modulation of T-bet desirable across a range of diseases.

Despite this theoretical attractiveness, transcription factors are inherently difficult to pharmacologically target due both to their intra-nuclear location, which precludes antibody based approaches, and also the inherent difficulty of targeting protein-DNA interactions given their large interaction domain. As such, transcription factors have been traditionally regarded as un-drugable (57). Given the difficulties of a small molecule approach, it is perhaps not surprising that most attempts at targeting T-bet therapeutically have centred on gene silencing approaches. By relying on the principle of base-pair complementarity and siRNA, such approaches bypass the need for complex protein or small-molecule engineering. In vivo, this approach has been used successfully in a range of pre-clinical models including in Experimental Allergic Encephalitis (EAE) (58). The Achilles heel of such genetic approaches is the difficulty of specifically targeting fragile oligonucleotides in vivo. This obstacle is mitigated by topical delivery around which most
Translational efforts have focused. Recently, a landmark Phase II study of inhaler delivered oligonucleotides targeting GATA-3 achieved primary end points, highlighting the potential of this approach (55). This was a significant breakthrough as a succession of monoclonal antibodies targeting numerous promising cytokines in asthma have all failed at the randomised control trial (RCT) stage, and thus this success underscores the therapeutic potential of targeting transcription factors directly. Following this early success, this approach targeting GATA-3 is currently being applied to Ulcerative Colitis in Phase II trials (ClinicalTrials.gov: NCT02129439). Separately, a T-bet inhibitor, based on the same gene silencing approach, is currently in pre-clinical development (59).

Alongside the development of gene silencing approaches to target transcription factors, the notion that transcription factors are un-drugable has been recently challenged by the development of small molecule inhibitors for the immune regulator, RORγT (60). This development, the result of high-throughput library screening followed by targeted chemical optimisation, has resulted in a number of compounds that inhibit RORγT activity both *in vitro* and *in vivo*, including an oral inhibitor. The lessons learnt from this approach have huge implications for therapeutics and basic transcription factor biology. In a study by Xiao et al, of three RORγT inhibitors developed, the most therapeutically effective, TMP778, moderately inhibited DNA-binding (61). In contrast, the most effective DNA-binding inhibitor in this study, TMP920, as determined empirically by ChIP-seq, was the least effective in the preclinical EAE autoimmune disease model - thus this data would seem to support the idea postulated by the network theory that it is protein-protein interactions that underscore transcription factor function. The rationale and catalyst, for the development of such inhibitors for RORγT was the view that Th17 cells are the most dominant cell in autoimmunity and thus the potentially large market for such
a therapy. The recent widespread consensus that pathogenic Th17 cells actually co-
express and require T-bet for pathogenicity, further underscores the therapeutic potential
of targeting T-bet in this manner (31,62,63).
Conclusion and unanswered questions in T-bet biology.

In summary, we have focused on the emerging role of T-bet in the mucosa and in particular used the highly idiosyncratic effects of T-bet on DCs and IELs, that are mirror opposites of T-bet’s standard effects on helper T-cells, to argue that structural determination of T-bet genomic binding through a conserved T-box recognition motif is insufficient to account for its increasingly diverse molecular effects. We have used evidence from ChIP-seq and evolutionary biology, and indeed the most extensive studies of transcription factor binding, to argue that the cell specific effects of T-bet are an output of the network of transcription factors existing in a cell prior to T-bet induction, thus providing an intuitively attractive explanation as to why T-bet induces unique effector programs in each of the cells it regulates. This argument, while strongly supported by existing evidence, is theoretical and requires further studies in the context of the novel populations regulated by T-bet.

This network theory is essentially an expansion of the long held view that negative gene regulation by T-bet is driven by protein-protein interactions - we have expanded this notion further to argue that protein-protein and not protein-DNA interactions are the key mediator of positive gene regulation by a transcription factor. Such a model, remains to be further validated in the context of T-bet.

Remarkably, despite 15 years of an exponential growth in our knowledge of the molecular effects of T-bet, our understanding of T-bet biology outside helper T-cells is in its infancy. The sum of our knowledge of T-bet’s genomic binding co-ordinates arises through ChIP-chip and ChIP-seq studies on helper T-cells alone. The development of more efficient techniques such as single cell ChIP-seq promise to expand this knowledge to much rarer populations including innate immune cells and DCs that have not been
amenable to traditional ChIP-seq assays (64). Furthermore, despite the existence of several transcriptional data-set from various T-bet deficient cells, there has been no attempt to integrate these systematically to analyse which T-bet targets are specific or unique to any of the lineages regulated by T-bet.

The empirical data suggests, and indeed as our network theory argues, that there are large numbers of T-bet protein interactions that are yet to be determined. While we currently know of over a dozen direct protein interaction between T-bet and other factors as illustrated in Figure II, these have largely been determined on a single molecule hypothesis driven basis. Robust determination requires system wide proteomics to detect such interactions in a systematic un-biased manner, as it is likely that these underlie cell-specific transcriptional effects. Determination of these effects is also likely to have implications for therapeutics as the experience with small molecular targeting of RORγt suggests that these protein-protein interactions are the key desirable target. Further elucidation of these networks may provide a manner to specifically target T-bet in distinct lineages, through targeting of T-bet’s interactions with cell-specific interaction partners.

Remarkably despite the huge potential of evolutionary approaches to deduce regulatory networks through conservation, this approach has remained largely untapped in immunology. Such approaches with other tissue specific transcription factors have yielded huge insights and are likely to do the same in immunology. As such, our broadened understanding of T-bet’s functional roles is likely to expand from the classical reductionist paradigm into the era of systems biology. This increased understanding, coupled with success at translating this into effective therapeutics may eventually herald the long-awaited translation of our knowledge of the canonical role of T-bet on immune responses into concrete therapeutics.
Acknowledgements

This study was supported by grants awarded by the Wellcome Trust (GL, grant number 091009) and the Medical Research Council (GL, grant number MR/M003493/1). RM is supported by a Wellcome Trust Clinical Training Fellowship (grant number 107345). Research was also supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St Thomas and King's College London (GL). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.
References


(9) Zheng W, Flavell RA. The Transcription Factor GATA-3 Is Necessary and Sufficient for Th2 Cytokine Gene Expression in CD4 T Cells Cell 1997;89(4):587 <last_page> 596.


(42) CHANG F, MAHADEVA U, DEERE H. Pathological and clinical significance of increased intraepithelial lymphocytes (IELs) in small bowel mucosa. APMIS 2005;113(6):385-399.


(59) Serdar Sel HR. Patent CA2776997A1 - Method for the production of a cell and/or
tissue and/or disease phase ... - Google Patents Available at:

(60) Huh JR, Littman DR. Small molecule inhibitors of RORgammat: targeting Th17
cells and other applications Eur J Immunol 2012 Sep;42(9):2232-2237.

RORgammat antagonists inhibit T helper 17 cell transcriptional network by divergent

(62) Harbour SN, Maynard CL, Zindl CL, Schoeb TR, Weaver CT. Th17 cells give rise
to Th1 cells that are required for the pathogenesis of colitis Proc Natl Acad Sci U S A
2015 Jun 2;112(22):7061-7066.

(63) Hirota K, Duarte JH, Veldhoen M, Hornsby E, Li Y, Cua DJ, et al. Fate mapping of
IL-17-producing T cells in inflammatory responses Nat Immunol 2011 Mar;12(3):255-
263.

(64) Rotem , Assaf, Ram , Oren, Shoresh , Noam, Sperling R.A., Goren , Alon, Weitz D.A.,
et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state.

global view on proteins and their functional interactions in 630 organisms. Nucleic Acids
Research 2009 January 01;37(suppl 1):D412-D416.
FIGURE LEGENDS

**Figure I: The pathogenic pathways implicated in TRUC colitis model.**
The first detectable perturbation in TRUC colitis is the increased activation of CD11C+ MHCII+ CD103- DCs that produce increased TNF in the absence of T-bet. TNF leads to epithelial cell apoptosis and compromise of the epithelial membrane integrity – this initiates colitis and is accompanied by the outgrowth of pathogenic microbiota. Initially disease is driven solely by TNF, though after several weeks, CD11C+ MHCII+ CD103- DC produced IL23 and TNF activate ILC3s. ILC3s produce inflammatory cytokines, including IL-17 which synergises with TNF to recruit neutrophils to colonic lamina propria. Thus while TNF is key in early stages of pathogenesis, later stages are driven by ILC activation and neutrophil recruitment.

Contributing to pathology is the selective outgrowth of gram negative pathogenic microbiota induced by an undefined deficiency in the innate immune system (A. and B). This outgrowth of pathogenic microbial species is detected by mucosal DCs, which produce IL-23 and increased levels of TNF in the absence of T-bet (C.), leading to epithelial cell apoptosis and increased ILC3 activation (D). Activated ILC3s produce IL-17 and IL-22 leading to further mucosal pathology from neutrophil recruitment and activation.

**Figure II: T-bet (TBX21) interaction partners.**
STRING interaction graph showing T-bet’s (TBX21) known experimentally determined protein –protein interaction partners in *Mus Musculus* (65)
Disclosure

The authors declare no conflict of interest.