Development 133, 1311-1322 (2006) doi:10.1242/dev.02305

A tyrosine-rich domain within homeodomain transcription factor Nkx2-5 is an essential element in the early cardiac transcriptional regulatory machinery

David A. Elliott[1,*,†], Mark J. Solloway[1,*,‡], Natalie Wise[1], Christine Biben[1], Mauro W. Costa[1,3], Milena B. Furtado[1], Martin Lange[1,§], Sally Dunwoodie[1,2] and Richard P. Harvey[1,2,¶]

Homeodomain factor Nkx2-5 is a central component of the transcription factor network that guides cardiac development; in humans, mutations in *NKX2.5* lead to congenital heart disease (CHD). We have genetically defined a novel conserved tyrosine-rich domain (YRD) within Nkx2-5 that has co-evolved with its homeodomain. Mutation of the YRD did not affect DNA binding and only slightly diminished transcriptional activity of Nkx2-5 in a context-specific manner in vitro. However, the YRD was absolutely essential for the function of Nkx2-5 in cardiogenesis during ES cell differentiation and in the developing embryo. Furthermore, heterozygous mutation of all nine tyrosines to alanine created an allele with a strong dominant-negative-like activity in vivo: ES cell↔embryo chimaeras bearing the heterozygous mutation died before term with cardiac malformations similar to the more severe anomalies seen in *NKX2.5* mutant families. These studies suggest a functional interdependence between the NK2 class homeodomain and YRD in cardiac development and evolution, and establish a new model for analysis of Nkx2-5 function in CHD.

KEY WORDS: Heart, Homeodomain, Nkx2-5, Congenital heart disease

INTRODUCTION

Genetic studies over the past decade have defined a cardiac regulatory network that is used for heart formation in both vertebrates and invertebrates (Cripps and Olson, 2002; Harvey, 1996). This network probably evolved for specification of a visceral muscle type specialised for pulsatile contraction (Harvey, 1996), and the cardiac-like contractile properties of body wall muscles in primitive diploblastic metazoa have long been appreciated (Fye, 1987).

Nkx2-5/Csx is a vertebrate member of the NK2 class of homeodomain transcription factors that sits high in the cardiac regulatory hierarchy (Cripps and Olson, 2002). Murine *Nkx2-5* was first identified in screens for relatives of the *Drosophila* gene *tinman*, expressed immediately downstream of mesodermal specification genes during fly development and absolutely required for formation of precursor cells of the heart and gut muscle lineages (Bodmer, 1993; Azpiazu and Frasch, 1993). Numerous vertebrate cognates of *Nkx2-5* have now been described, with related genes expressed in the simple heart tube of amphioxus (Holland et al., 2003), in the pulsatile muscular pharynx of *C. elegans* (Okkema et al., 1997) and in the contractile foot of hydra (Shimizu and Fujisawa, 2003). Transgenic rescue experiments testify to the functional homology that exists between cognates of this gene family in distantly related species (Haun et al., 1998; Park et al., 1998; Ranganayakulu et al., 1998).

Nkx2-5 is expressed in the earliest recognisable cardiac precursor cells in all models examined (Harvey, 1996), including cells of the heart morphoregulatory field in *Xenopus* (Raffin et al., 2000), and the second heart precursor field of amniotes (Stanley et al., 2002). In *Xenopus*, dominant-negative inhibition of Nkx2-5 and its close relative Nkx2-3 leads to total loss of all cardiac progenitors (Grow and Krieg, 1998). In homozygous *Nkx2-5* knockout mouse embryos, a simple beating myogenic heart tube is able to form, although differentiation and morphogenesis of specialised chamber myocardium is blocked (Lyons et al., 1995). A remarkable diversity of structural and functional abnormalities of the heart occur in human families and individuals carrying heterozygous *NKX2.5* mutations (Benson et al., 1999). Atrioventricular (AV) conduction block, present in most individuals, results from hypoplasia and progressive postnatal loss of AV nodal tissue, and requires a pacemaker (Jay et al., 2004; Pashmforoush et al., 2004). Structural malformations requiring surgery include atrial septal defect (ASD), tetralogy of Fallot, ventricular septal defect (VSD) and hypoplastic left heart syndrome (Benson et al., 1999; Elliott et al., 2003; McElhinney et al., 2003).

Among the earliest Nkx2-5-dependent genes are those encoding other cardiac transcription factors (Biben and Harvey, 1997; Bruneau et al., 2000; Molkentin et al., 2000; Ueyama et al., 2003; von Both et al., 2004; Yoshioka et al., 1998). Nkx2-5 physically interacts with several transcription factors, including members of the GATA and T-box factor families, serum response factor (SRF) and Foxh1, and collaborates with them to regulate target promoters (Chen and Schwartz, 1996; Durocher et al., 1997; Bruneau et al., 2001; Hiroi et al., 2001; von Both et al., 2004). Recent data show that competition between Nkx2-5 and different members of the T-box family of transcription factors drives formation of chamber and non-chamber myocardium (Habets et al., 2002; Stennard and Harvey, 2005). Nkx2-5

[1]Victor Chang Cardiac Research Institute, Darlinghurst, Sydney 2010, Australia. [2]Faculties of Life Science and Medicine, University of New South Wales, Randwick 2031, Australia. [3]Instituto de Biofisica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, 20941-000, Brazil.

*These authors contributed equally to this work
[†]Present address: The Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK
[‡]Present address: Genentech, Molecular Biology Department, 1 DNA Way, South San Francisco, CA 94080-4990, USA
[§]Present address: Max Planck Institute for Molecular Genetics, Cardiovascular Genetics, Ihnestrasse 73, 14195 Berlin, Germany
[¶]Author for correspondence (e-mail: r.harvey@victorchang.unsw.edu.au)

Accepted 31 January 2006

also activates negative regulatory circuits controlled by transcriptional repressors CARP and HOP (Chen et al., 2002; Zou et al., 1997).

Despite these advances, the central question of how broadly expressed transcription factors, such as Nkx2-5, control the temporal and spatial specificity of lineage and morphogenetic events during heart development remains unanswered. Region-specific signalling inputs that influence the nuclear localisation and/or the chromatin-modifying activity of transcription factors are likely to play key roles (Charron et al., 2001; McKinsey et al., 2000). Members of the cardiac Nkx2 family are known to function as both transcriptional activators and repressors (Choi et al., 1999), and two conserved domains, the N-terminal TN-Domain and C-terminal NK2-specific domain (NK2SD) appear to act as negative modulatory domains, the former through association with the Groucho family of co-repressors (Muhr et al., 2001; Watada et al., 2000).

In this paper, we genetically define a novel, conserved and essential transcriptional domain within the C terminus of murine Nkx2-5. The tyrosine-rich domain (YRD) has co-existed with the NK2 class homeodomain since before radiation of vertebrate and invertebrate evolutionary lines. The apparent absolute functional inter-dependence between the YRD and NK2 homeodomains can explain the similar phenotypic manifestations arising from mutations in different regions of the NKX2.5 protein in individuals with CHD.

MATERIALS AND METHODS

Expression constructs and culture assays

Plasmids for G5E1b-luciferase, pGAL, FLAG-Nkx2-5, 3xHA-luciferase and *Nppa1* and *Gja5* promoters, and transient transfection protocols were as described previously (Ranganayakulu et al., 1998; Sadowski et al., 1992; Stennard et al., 2003). Neonatal rat cardiomyocytes were cultured (Engelmann, 1993) and transfected using Lipofectamine and Lipofectamine PLUS (GIBCO BRL).

Western blotting and EMSA

Nuclear extracts were prepared (Stennard et al., 2003) and proteins separated by SDS-PAGE. Western blotting was performed using anti-FLAG mouse mAb (1:1000; AMRAD) or anti-GAL4 (DBD) mouse mAb (RK5C1) (1:1000; Santa Cruz) with anti-mouse IgG-HRP (1:2000; Silenus). Proteins were detected using ECL reagent (Pharmacia Amersham Biotech). Nkx2-5 proteins for EMSA (Stennard et al., 2003) were translated using TNT Rabbit reticulate lysates (Promega). The oligonucleotide contained three high-affinity Nkx2-5-binding sites (5′ ctcaagtgg 3′).

Nkx2-5 homologues

Systemic or branchial hearts from adult *O. kaurna*, *S. officinalis* and *L. pealeii* were dissected into RNAlater (Ambion). RNA was extracted using Trizol (Invitrogen). RT-PCR was performed using 200 ng of NK2 homeodomain-specific degenerate oligonucleotides (5′ cgictittytcicaigc 3′ and 5′ rtaicgicgittgtgiaacc 3′; r=a or g; i=deoxyinosine), 0.5 mM dNTPs, 1×PCR buffer (Boehringer) and 2U of TAQ (Boehringer). Cycling conditions were: 95°C for 2 minutes followed by 35 cycles of 95°C for 20 seconds, 50°C for 30 seconds and 72°C for 1 minute. Homeobox sequences were used to design specific primers for 5′ and 3′ RACE using Ambion and Clontech kits, respectively, and a modified oligonucleotide (5′ aagcagtggtatcaacgcagagtacgcgggtttt 3′).

ES cells and gene targeting

Gene targeting was performed using standard methods. The IRESlacZHygro cassette was adapted from IRESCreHygro (Stanley et al., 2002). Embryoid body (EB) culture was adapted from the hanging drop method (Bader et al., 2001). GFP expression was analysed on a FACSCalibur (BD Biosciences). A sort gate was established on the basis of forward scatter, side scatter and propidium iodide staining. A second gate was established based on GFP fluorescence intensity (FL-I) and side scatter that excluded most auto-fluorescent cells. Animal experiments were approved by the St Vincent's Hospital/Garvan Institute for Medical Research Animal Ethics Committee.

RT-PCR analysis

RNA from FACS-isolated GFP-positive cells was extracted using Trizol (Invitrogen). After RNA extraction, cDNA was prepared and linearly amplified (Baugh et al., 2001). Results were confirmed by RT-PCR without amplification. Real-time PCR used the LightCycler-FastStart DNA master SYBR Green I kit (Roche) on a LightCycler (Roche). PCR protocols: 95°C for 10 minutes; followed by 35 cycles of 95°C for 10 seconds; 60°C for 10 seconds; 72°C for 10 seconds.

TUNEL assay

TUNEL was performed on sections from three 14.5 dpc control *Nkx2-5$^{lacZ/+(\Delta H)}$*↔wild-type and two similarly staged *Nkx2-5$^{Y\text{-}A:IRESlacZ/+(\Delta H)}$*↔wild-type chimaeric hearts using the DeadEnd fluorometric TUNEL system (Promega). TUNEL was assessed on a total of 8-16 sections from the mid-region of chimaeric hearts. Statistical significance was assessed using nested analysis of variance.

RESULTS

A tyrosine-rich domain in the C-terminus of Nkx2-5 is required for its transcriptional activity in vitro

We have previously shown that C-terminal amino acids 232-318 of murine Nkx2-5 possess positive transcriptional activity in a heterologous assay in C3H10T1/2 fibroblasts (Ranganayakulu et al., 1998). This activity does not require the conserved NK2SD implicated in transcriptional repression. To define the region(s) essential for positive activity, the C-terminal fragment and progressively smaller sub-fragments were fused to the yeast GAL4 DNA-binding domain (DBD) and tested for transcriptional activity on a GAL4-dependent luciferase reporter gene (G5T-luc) in C3H10T1/2 cells (Fig. 1A,B). Only the full length fragment [Nkx(232-318)] showed activity, with deletion of 30 or more N-terminal amino acids reducing activity to background levels. This finding suggests that the conserved C-terminal GIRAW motif and Nkx2-5 Box (Evans, 1999), the functions of which are unknown, have no autonomous transcriptional activity in this assay. To confirm this finding, we tested whether the GIRAW motif and Nkx2-5 Box were essential in the context of the full-length C-terminal fragment. Mutation of GIRAW to GIRAG, GIGAW or merely GI failed to diminish transcriptional activity significantly (Fig. 1C). An alanine scanning mutation series was constructed across the Nkx2-5 Box and the central VGDL peptide was also deleted (Fig. 1D). Most of these mutations had little or no effect on transcription. Others, notably D293A and ΔVGDL, reduced activity to 44-48% of control levels, although this minor effect may be due to the slightly diminished stability of mutant proteins (Fig. 1F).

We noted that the region essential for transcriptional activity, minimally encompassing amino acids 232-262, was unusually rich in the bulky aromatic amino acid tyrosine (24% over 37 amino acids). To assess whether tyrosines were essential for transcriptional activity, we first mutated each individually to alanine (Fig. 1E). Individual Y→A mutations had only minor effects on transcription, in the extreme diminishing activity by 29% in the case of Y244A and increasing it by 52% in the case of Y363A (Fig. 1E). However, mutation of the first seven out of nine tyrosines to alanine as a group reduced activity to 25% of wild type, and when all nine were mutated, activity was reduced further. Mutation of the first seven tyrosines to the structurally related phenylalanine also effectively eliminated activity. Relevant mutant proteins were stable after transfection of encoding plasmids into COS cells (Fig. 1F; data not shown).

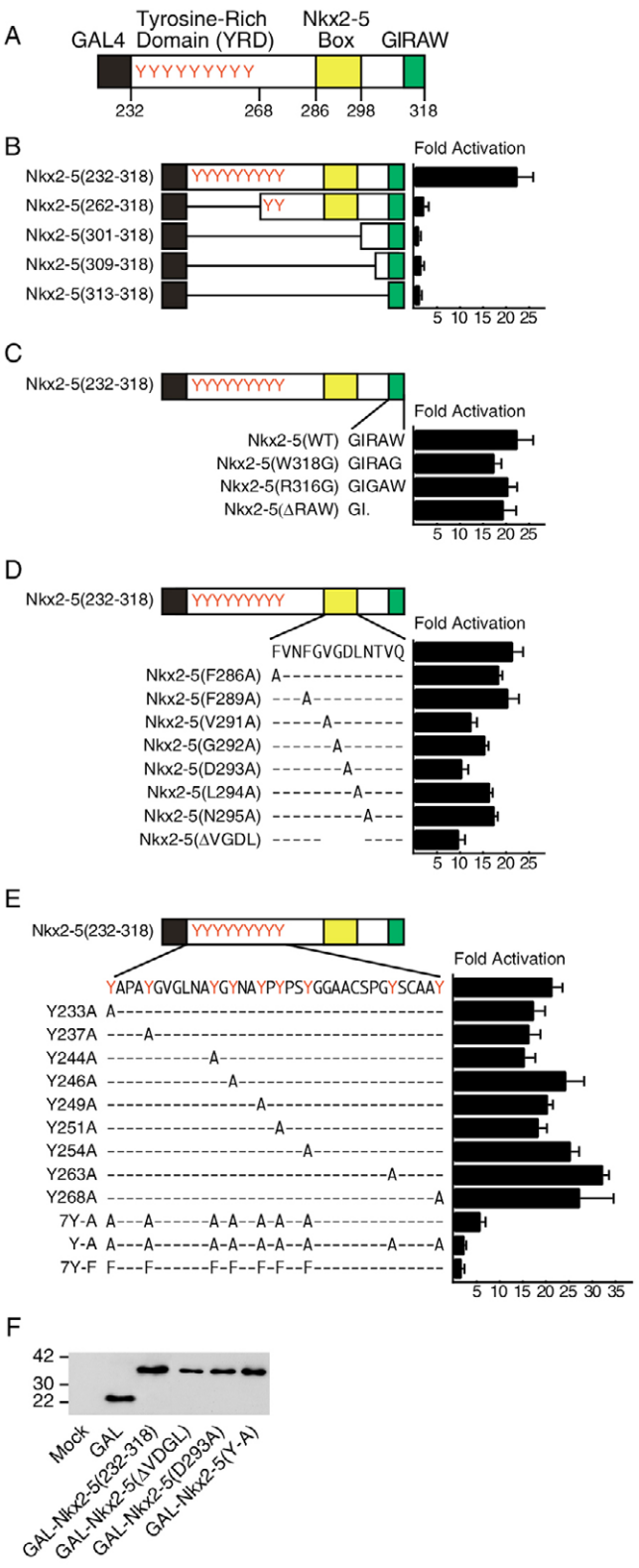

Fig. 1. Transcriptional activity in the C terminus of Nkx2-5. (**A**) Domain structure of the Nkx2-5 C terminus fused to GAL4 DNA-binding domain (GAL4). (**B-E**) Trans-activation of the GAL4-dependent G5T-luciferase reporter in C3H10T1/2 cells by co-transfected GAL-Nkx2-5(232-318) and similar constructs bearing mutations (see text). Relative transcriptional activity compared with vector-only control (bars indicate s.e.m.) is shown on the right of each panel. (**F**) Western blot showing that GAL-Nkx2-5 proteins are stably expressed in transfected COS cells.

Conservation of the YRD in the cardiac clade of NK2 homeodomain proteins

A YRD was evident C-terminal to the homeodomain and NK2SD in all known chordate members of the 'cardiac' subclass of NK2-class homeodomain proteins (Harvey, 1996), including a member from the cephalochordate amphioxus (amphiNk2-tin) (Fig. 2A). These homeodomain proteins form an evolutionary clade based on sequence similarities within their homeodomains and because of common expression in developing cardiac and/or visceral muscles. A tyrosine-rich element is also present in *Drosophila* bagpipe (bap) (Azpiazu and Frasch, 1993) and its vertebrate homologues [i.e. Nkx2-5 relatives that are expressed in gut mesoderm and other organ and axial systems (Fig. 2A; see Fig. S3 in the supplementary material)]. The YRD was absent from members of a sister clade to the cardiac Nkx2 proteins, including Nkx2-1, Nkx2-2 and Nkx2-4, that are expressed in the nervous system.

Alignment of YRD sequences (Fig. 2A) revealed that the number of tyrosines varied from four to 10 in different NK2 proteins, although there was an overall conservation of spacing, particularly among species orthologues of Nkx2-5 and Nkx2-3. The alignment became more significant if the occasional conservative amino acid change to phenylalanine was allowed, supporting the idea that the tyrosines and phenylalanines are structurally or functionally important. In addition to tyrosines, several other amino acids were strongly, although not absolutely, conserved. The asparagine at position 15 in mouse Nkx2-5, for example, was conserved in all of the Nkx2-5 and Nkx2-3 orthologues shown (Fig. 2A) and valine at position 7 was conserved in 10/11 of these. Asparagine 10, prolines 11, 18 and 37, and cysteine 33 in the mouse sequence were also well conserved in other members.

Cardiac NK2 proteins from cephalopod molluscs also carry the YRD

The YRD was absent from *Drosophila* Tinman, despite it having a well-established cardiogenic function during *Drosophila* embryogenesis (Cripps and Olson, 2002). A cluster of tyrosines was present C-terminal to the homeodomain in the *C. elegans* NK2 homeoprotein, Ceh22, which has a role in specification of the pulsatile (heart-like) pharyngeal muscles of the worm, a function that can be substituted for by Tinman (Haun et al., 1998). However, other signature amino acids of the YRD are not conserved. Furthermore, the NK2SD is lacking in Ceh22. Both insect and nematode phyla are members of the large invertebrate clade Ecdysozoa. Accelerated evolution of gene families has been noted in the genomes of *Drosophila* and *C. elegans* (Kortschak et al., 2003), raising the possibility that the absence of C-terminal domains in Tinman and Ceh22 represents a derived rather than ancestral state.

To explore this possibility, we cloned cDNAs encoding *Nkx2-5* homologues from the hearts of cephalopod molluscs (octopus *Octopus kaurna*, cuttlefish *Sepia officinalis* and squid *Loligo pealii*). Cephalopod molluscs belong to a separate large invertebrate clade, Lophotrochozoa. They are highly motile invertebrates that possess a sophisticated closed circulatory system with systemic and branchial hearts (Fig. 2B). Remarkably, the systemic hearts of squid produce pressures approaching those of mammalian hearts (Wells, 1992). A full-length cDNA clone was isolated from RNA extracted from the adult systemic heart of the cuttlefish *S. officinalis*, and shorter clones were isolated from *O. kaurna* and *L. pealii*. In situ hybridisation to sectioned *L. pealii* embryos confirmed expression in the systemic heart and adjacent muscular ink sac (Fig. 2B-D; data not shown).

The cuttlefish *SoNkx2-5* cDNA was 1.12 kb long and encoded a 326 amino acid protein containing, in addition to an NK2-class homeodomain, four other conserved NK2 protein signature motifs: a TN Domain within the N terminus and (within the C terminus) the NK2SD, YRD and GIRAW motifs (Fig. 2A). The same C-terminal domain structure was evident in proteins predicted from the *O. kaurna* and *L. pealii* cDNAs (data not shown). The Nkx2-5 Box, found in only vertebrate and amphioxus Nkx2-5 orthologues (Evans, 1999), was absent.

The cephalopod homeodomain sequences carried a tyrosine at position 54, a defining hallmark of NK2-class homeodomains and essential for their unique DNA-binding site specificity (Tsao et al.,

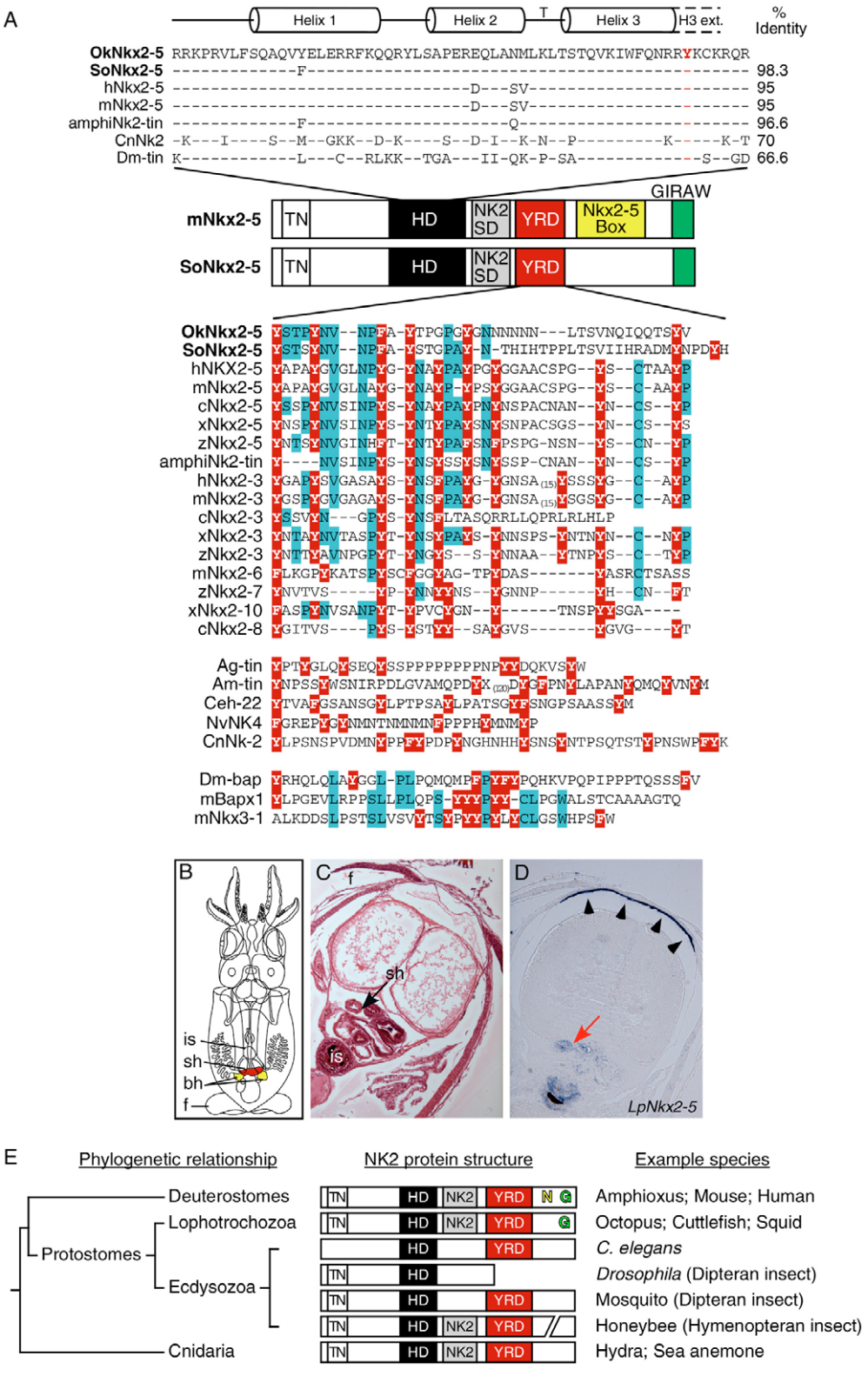

Fig. 2. Cephalopod mollusc cardiac NK2 proteins carry the YRD and show a domain structure resembling their mammalian relatives. (**A**) Homeodomain and YRD sequences from cephalopod and other cardiac NK2 proteins. Tyrosine 54 within the homeodomain, a defining hallmark of NK2-class homeodomains, is shown in red. Tyrosines and phenylalanines within the YRD are shown in red, with other conserved amino acids in green. Species abbreviations: Ag, *Anopheles gambiae* (mosquito); Am, *Apis mellifera* (honeybee); amphi, amphioxus; c, chick; Ce, *Caenorhabditis elegans* (nematode); Cn, cnidarian *Hydra vulgaris* (hydra); Dm, *Drosophila melanagaster*; h, human; m, mouse; Nv, *Nematostella vectensis* (sea anemone); Ok, *Octopus kaurna* (octopus); So, *Sepia officinalis* (cuttlefish); x, *Xenopus laevis* (frog); z, zebrafish. TN, Tin-domain. (B-D) *LpNkx2-5* expression in *Loligo paelii* (squid) embryos. (**B**) Schematic diagram of a squid embryo showing location of systemic heart (sh; red) and branchial hearts (bh; yellow). (**C**) Histological section showing systemic heart and ink sac (is). (**D**) In situ hybridisation to an adjacent section showing *LpNkx2-5* expression in the systemic heart and ink sac (arrow), as well as muscle beneath the fins (f; arrowheads). (**E**) Phylogenetic relationships and domain structure of NK2 protein from divergent species. G, GIRAW motif; N, Nkx2-5 box; NK2, NK2SD.

1994). Homeodomain comparisons clearly demonstrated that the cephalopod proteins were more closely related to their vertebrate Nkx2-5 and Nkx2-3 relatives than to Tinman (Fig. 2A). For example, the SoNkx2-5 homeodomain was 95% identical to that of mouse and human Nkx2-5, 96.6% identical to amphioxus Nk2-tin, but only 66.6% identical to Tinman.

The YRD of the cephalopod proteins contained five or six tyrosines with an additional conservative change to phenylalanine and several non-tyrosine amino acids present in mammalian family members were also conserved (Fig. 2A). The above data suggest that the domain architecture of vertebrate cardiac NK2 homeodomain proteins, as seen in Nkx2-5, was established prior to divergence of the common ancestor of protostomes and deuterostomes, and that Tinman is highly derived.

Mutation of the YRD alters Nkx2-5 transcriptional activity but not DNA binding in vitro

Full-length Nkx2-5 is a weak transcriptional activator in vitro and it has been proposed, based on in vitro data, that the C-terminal region of the protein inhibits the activity of its strong N-terminal transactivation domain via an intramolecular mechanism (Sepulveda et al., 1998). An influence of C-terminal amino acids on Nkx2-5 dimerisation on DNA has also been reported (Kasahara et al., 2000). We tested whether the Y-A mutation affected the function of full-length Nkx2-5 in vitro. However, we found no diminishment of DNA binding (Fig. 3A) and there was no change in the transactivation activity of Nkx2-5Y-A on a synthetic promoter carrying multiple Nkx2-5-binding sites (Fig. 3B), or on the promoter of the *Nppa1* gene, a direct Nkx2-5 target, in the absence or presence of cardiac transcription factors Gata4 or Gata5, and Tbx20, with which Nkx2-5 can functionally synergise (Fig. 3C and Fig. S1 in the supplementary material). Furthermore, no effect was seen on the *Nppa* promoter in the presence of SRF and/or myocardin, and Tbx5 and/or Tbx2 (see Fig. S1 in the supplementary material). However, Nkx2-5Y-A transcriptional activity was diminished by 50% ($P<0.001$) relative to wild-type Nkx2-5 on the promoter of the *Gja5* gene (encoding connexin 40), another direct target of Nkx2-5; this was most obvious in the presence of Tbx20a and Gata5 (see Stennard et al., 2003) (Fig. 3C).

Introduction of the Y-A mutation into the native Nkx2-5 locus

Evolutionary conservation of the YRD in vertebrate, cephalochordate and invertebrate species strongly suggests functional significance that may not be strongly revealed by available in vitro assays. We therefore introduced the Y-A mutant allele into the native *Nkx2-5* locus using gene targeting. Four separate targeting experiments were performed to create an allelic series for analysis of YRD function in vitro and in vivo (Fig. 4A-C). In targeting experiment I, we mutated one allele of *Nkx2-5* in wild-type ES cells by insertion of a gene expression cassette encoding a nuclear-localising version of *E. coli* β-galactosidase (*lacZ*) into the first coding exon creating a new null allele. To ensure that no abnormal DNA-binding protein was produced from this allele, the vector also carried a deletion within the homeodomain region (ΔHD; Fig. 4B), as previously described (Biben et al., 2000). In this and other targeting vectors, drug resistance gene cassettes (hygromycin or neomycin) were flanked by LoxP sequences, the recognition sites of bacteriophage Cre

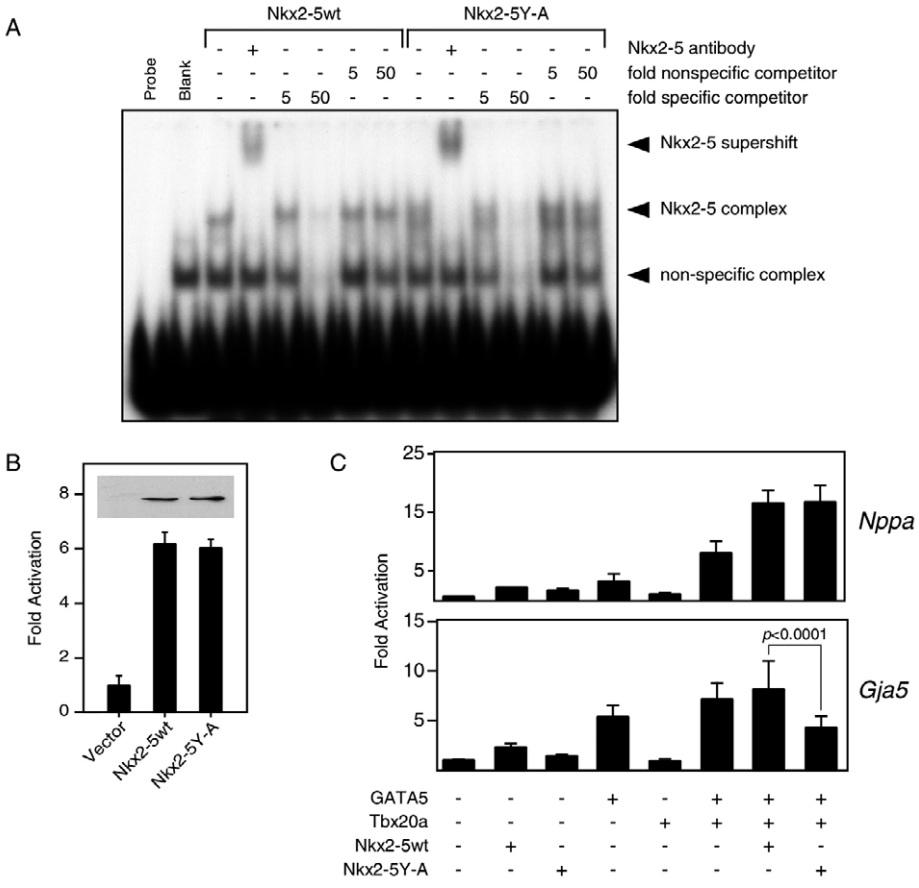

Fig. 3. DNA binding and transactivation analysis of the Nkx2-5Y-A mutant protein. (**A**) EMSA comparing DNA binding of in vitro-translated Nkx2-5 and Nkx2-5Y-A using an Nkx2-5-binding oligonucleotide probe. Anti-Nkx2-5 antibody and specific and non-specific competitor oligonucleotides were used to assess binding specificity. (**B**,**C**) In vitro transcriptional analysis using FLAG-tagged Nkx2-5 and Nkx2-5Y-A expression vectors and a multimerised Nkx2-5-binding site synthetic enhancer-luciferase reporter (B) or with and without Gata5 and Tbx20a on the *Nppa* (–700) and *Gja5* (–1190 to +121) promoter-luciferase reporters in C3H 10T1/2 cells (C). Inset in B shows western blot of transfected proteins. Significance assessed using Student's *t*-test.

recombinase. After validation of correctly targeted ES cell clones using Southern blotting (Fig. 4D), cassettes were removed by transient transfection of ES cells with a Cre recombinase expression vector. Following plating at clonal density, and clone selection and expansion, cassette deletion was validated using polymerase chain reaction (PCR) (alleles ΔH and ΔNH in Fig. 4B,C,E). Strategy I produced ES cell lines with genotypes *Nkx2-5*$^{lacZ/+}$ and *Nkx2-5*$^{lacZ/+(\Delta H)}$, both heterozygous null for *Nkx2-5* (Biben et al., 2000) and expressing β-galactosidase under *Nkx2-5* control.

In targeting strategy II (Fig. 4A), we used the same *lacZ* targeting vector to mutate the remaining wild-type allele of *Nkx2-5* in a previously established ES cell line (*Nkx2-5*$^{GFP/+}$) in which one allele of *Nkx2-5* had already been targeted by insertion of a gene cassette encoding an enhanced version of jellyfish green fluorescent protein (eGFP) (Biben et al., 2000). This strategy created ES cell lines with genotypes *Nkx2-5*$^{GFP/lacZ}$ and *Nkx2-5*$^{GFP/lacZ(\Delta NH)}$, homozygous null for *Nkx2-5* and expressing both eGFP and β-galactosidase under *Nkx2-5* control.

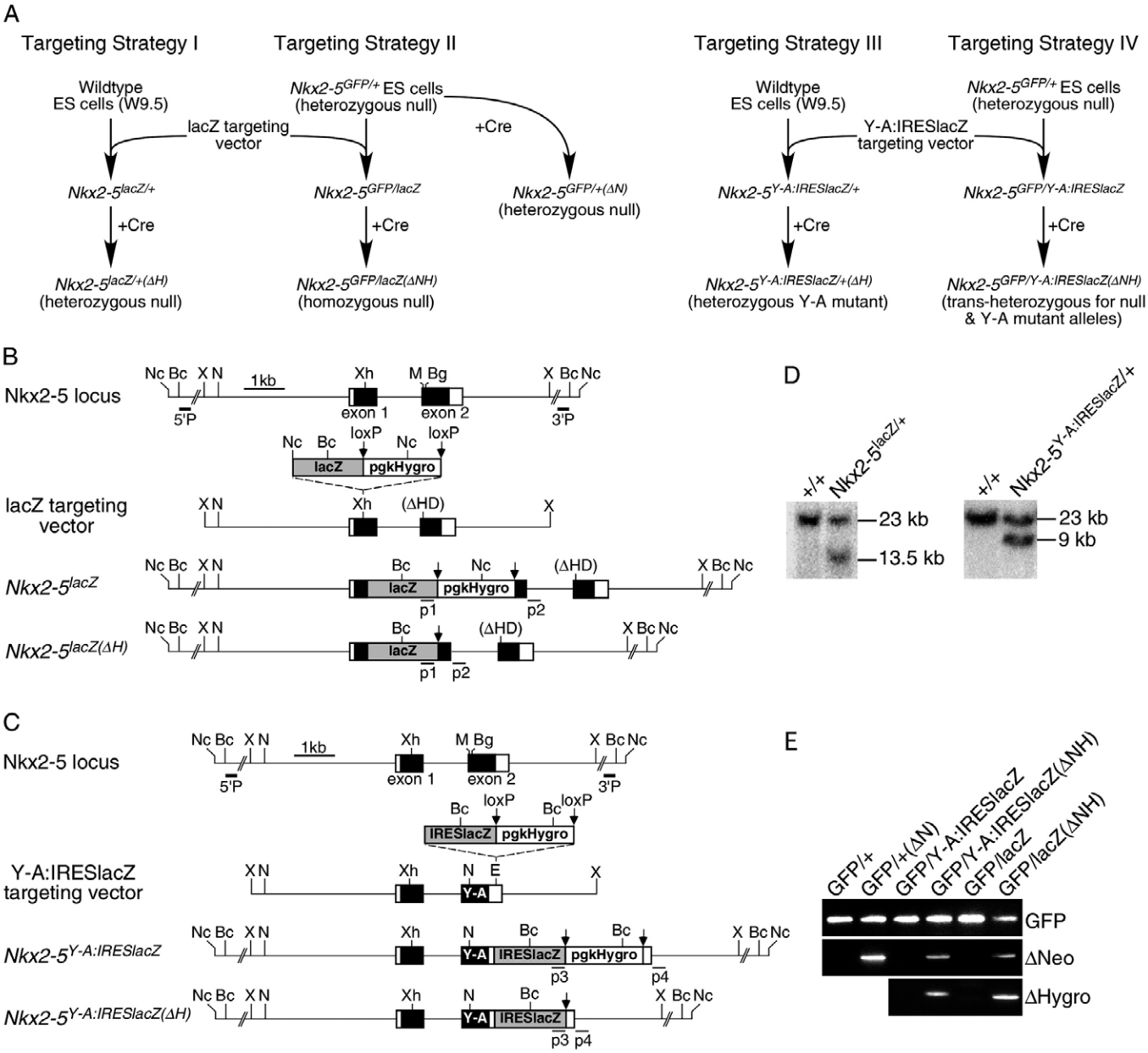

Fig. 4. Targeted alleles of *Nkx2-5*. (**A**) Targeting strategies used to create Nkx2-5 'knock-in' alleles (see text). (**B**,**C**) Wild-type *Nkx2-5* locus with targeting vectors and targeted alleles shown before and after Cre deletion of drug selection cassettes (H, pgkHygro; N, pgkNeo). Coding region exons shaded black. Arrows show loxP sites. ΔHD indicates the small homeobox deletion (amino acids 141-184). 5′P and 3′P, and p1-p4, respectively, indicate positions of Southern probes and PCR primers for validation of recombinations and deletions. Restriction endonucleases: Nc, *Nco*I; Bc, *Bcl*I; Bg, *Bgl*II; M, *Mlu*I; N, *Not*I; X, *Xba*I; Xh, *Xho*I. (**D**) Southern blot analysis of ES cell DNA of indicated genotype digested with *Bcl*I (left) or *Nco*I (right) using probe 3′P. The 23 kb band is the wild-type allele, while 9 and 6 kb bands represent targeted alleles. (**E**) PCR validation of Cre-deleted ES cell clones using primers specific for the pgkHygro cassette, as indicated in B and C, and those specific for the GFP and pgkNeo cassettes.

In strategy III, we targeted one *Nkx2-5* allele in wild-type ES cells using a vector (Fig. 4C) carrying a mutated YRD domain in which all nine tyrosines were mutated to alanine (Y-A mutation; Fig. 1E), as well as a picornavirus internal ribosome entry site (IRES)-*lacZ* gene cassette inserted into the 3′ untranslated region (Stanley et al., 2002). This strategy created ES cells with genotypes *Nkx2-5*$^{Y\text{-}A:IRESlacZ/+}$ and *Nkx2-5*$^{Y\text{-}A:IRESlacZ/+(\Delta H)}$, heterozygous for the Y-A allele and expressing β-galactosidase under *Nkx2-5* cis-regulatory control via IRES-mediated translational initiation of a bicistronic mRNA.

Finally, in strategy IV, we targeted the remaining wild-type *Nkx2-5* allele in *Nkx2-5*$^{GFP/+}$ ES cells with the Y-A vector, creating ES cell lines with genotypes *Nkx2-5*$^{GFP/Y\text{-}A:IRESlacZ}$ and *Nkx2-5*$^{GFP/Y\text{-}A:IRESlacZ(\Delta NH)}$, expressing one Y-A allele in the context of the null GFP allele, and both eGFP and β-galactosidase under *Nkx2-5* control.

To confirm correct marker expression in targeted clones, ES cells carrying mutant genotypes were differentiated in vitro into embryoid bodies (EBs) until day 5, then allowed to adhere to collagen-coated slides until day 9. Foci of beating cardiomyocytes were evident in 90% of colonies. In both *Nkx2-5*$^{GFP/lacZ(\Delta NH)}$ and *Nkx2-5*$^{GFP/Y\text{-}A:IRESlacZ(\Delta NH)}$ EB colonies, eGFP fluorescence and *lacZ* staining were coincident with each other, and with foci of beating cardiomyocytes (see Fig. S2A-D in the supplementary material; data not shown). We also confirmed stable expression of the Nkx2-5Y-A mutant protein using the EB system (see Fig. S2E-J in the supplementary material).

Cardiomyocyte gene expression is altered in *Nkx2-5*$^{GFP/Y\text{-}A:IRESlacZ}$ embryoid bodies

Cardiomyocytes formed within differentiating EBs represent cells of different states of chamber maturation and conduction phenotypes (Fijnvandraat et al., 2003). To compare gene expression in eGFP-positive cardiocytes of different Nkx2-5 genotypes, we developed a protocol for purifying Nkx2-5-GFP-positive cells from EBs using fluorescence-activated cell sorting (FACS) for eGFP fluorescence (Fig. 5A). Contamination with eGFP-negative cells was only 0.1-2% using this protocol. Virtually all surviving cells (>99%) were beating after replating and culture for 24 hours, suggesting that contamination with other cell types capable of expressing Nkx2-5, including foregut endoderm (Stanley et al., 2002), was minimal. eGFP-positive cells null for Nkx2-5 (*Nkx2-5*$^{GFP/lacZ(\Delta NH)}$) and those expressing only the Y-A mutant allele (*Nkx2-5*$^{GFP/Y\text{-}A:IRESlacZ(\Delta NH)}$) showed similar fluorescence versus cell count profiles, but these profiles were qualitatively different from that of cells heterozygous for the null eGFP allele (*Nkx2-5*$^{GFP/+(\Delta N)}$) (Fig. 5B).

To explore this difference further, we analysed expression of a subset of Nkx2-5-dependent genes in FACS-purified EB-derived cardiomyocytes of different genotypes, using comparative reverse transcriptase (RT)-PCR and quantitative real-time RT-PCR (Fig. 5C,D). Genes known to be downregulated in the context of *Nkx2-5* mutation or dominant-negative inhibition in vivo were analysed, including transcription factors *Hand1*, *Irx4* and *Ankrd1* (encoding CARP) as well as *Myl2* (encoding myosin light chain 2v), *Nppa1* (encoding atrial natriuretic factor), *Tagln* (encoding sm22α), *Smpx* (encoding Chisel) and *Gja1* (encoding connexin 43). Normalised to levels of *gapd* expression, all of these targets, with the exception of *Hand1*, were downregulated significantly and to the same extent in cardiomyocytes of the null (*Nkx2-5*$^{GFP/lacZ(\Delta NH)}$) and YRD-over-null (*Nkx2-5*$^{GFP/Y\text{-}A:IRESlacZ(\Delta NH)}$) mutant genotypes, relative to levels in FACS-purified heterozygous null *Nkx2-5* mutant cells (*Nkx2-5*$^{GFP/+(\Delta N)}$). *Hand1* was only moderately affected in null and Y-A mutant cells, suggesting that its downregulation in *Nkx2-5* mutant hearts (Biben and Harvey, 1997) is indirect.

Analysis of YRD function in chimaeric embryos

To examine the phenotypic effects of the YRD mutation in whole embryos, we created chimaeric embryos by injection of ES cells carrying *Nkx2-5*$^{lacZ/+(\Delta N)}$, *Nkx2-5*$^{GFP/lacZ(\Delta NH)}$ and *Nkx2-5*$^{GFP/Y\text{-}A:IRESlacZ(\Delta NH)}$ genotypes into wild-type blastocysts. Injected

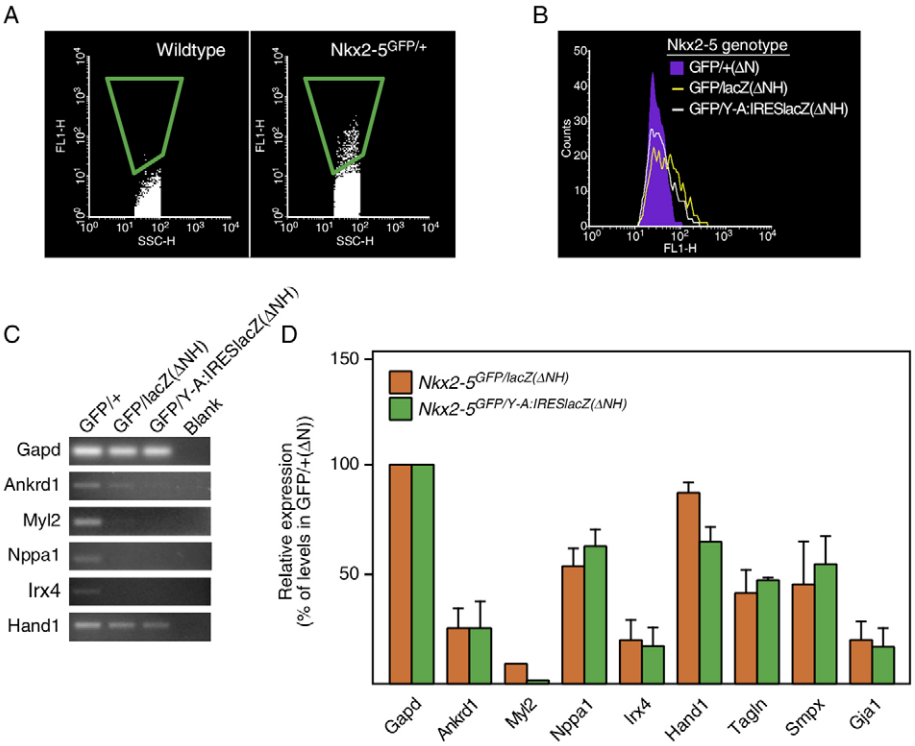

Fig. 5. The Nkx2-5Y-A allele impairs cardiac differentiation in EBs. (**A**) FACS profiles of dissociated wild-type and *Nkx2-5*$^{GFP/+}$ EBs showing side scatter (SSC-H) versus GFP fluorescence (FL1-H). Gated cells were collected for expression profiling. (**B**) Comparison of GFP fluorescence versus cell counts in EBs of different genotype. (**C**) Comparative RT-PCR analysis of gene expression in FACS-purified GFP+ cells from EBs of genotypes indicated. (**D**) Real-time RT-PCR analysis of gene expression from FACS-purified GFP+ cells as in C. Results reported as percent expression relative to *Nkx2-5*$^{GFP/+(\Delta N)}$ cells (two independent experiments performed in duplicate; error bars indicate ±s.e.m.).

blastocysts were surgically transferred to pseudo-pregnant mothers for fostering, and embryos were subsequently harvested at different stages of development for *lacZ* staining and phenotypic analysis. For all genotypes, *lacZ* was expressed with the correct pattern in the cardiac crescent at ~7.75 dpc and looping heart tube at ~8.0 dpc (Fig. 6A,B) (Stanley et al., 2002). In *Nkx2-5*$^{lacZ/+(\Delta N)}$↔wild-type chimaeras in which there was a high contribution of ES cell progeny (judged by the extent of *lacZ* staining in the heart; Fig. 6D,G), the heart looped normally at 9.0-9.5 dpc and heterozygous cells had contributed to chamber and non-chamber myocardium, trabeculae and extra-cardiac regions. Furthermore, these chimaeras transmitted the mutant allele through the germline when taken to term. By contrast, high level *Nkx2-5*$^{GFP/lacZ(\Delta NH)}$↔wild-type and *Nkx2-5*$^{GFP/Y\text{-}A:IRESlacZ(\Delta NH)}$↔wild-type chimaeras arrested development at ~8.5 dpc, showing pericardial oedema, failure of cardiac looping, lack of ventricular chamber discrimination, no trabeculation and a truncated outflow tract (Fig. 6C,E,F,H,I), a close phenocopy of *Nkx2-5*-null mutant hearts (Lyons et al., 1995).

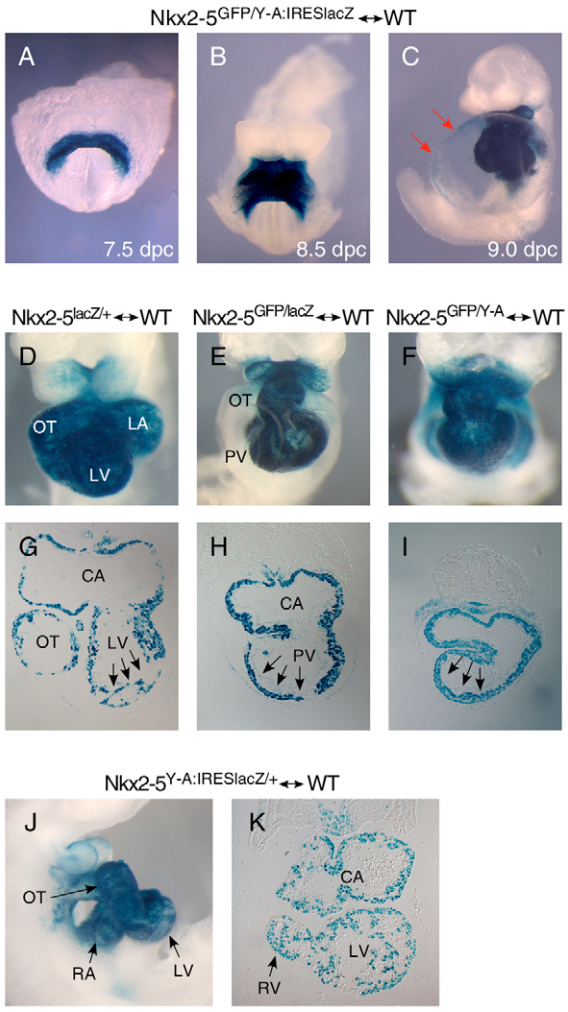

Fig. 6. High-level *Nkx2-5*$^{GFP/lacZ}$ or *Nkx2-5*$^{GFP/Y\text{-}A}$ ES cell contribution to chimaeric embryos perturbs heart development. (**A**,**B**) *Nkx2-5*$^{GFP/Y\text{-}A}$↔wild-type chimaeras at 7.5 and 8.5 dpc stained for *lacZ*. (**C**) Chimaera showing abnormal cardiac development at 9.0 dpc with pericardial oedema (arrows). (**D-I**) Comparison of heart morphology (upper panels) and histology (lower panels) in 9.0 dpc chimaeric embryos composed of indicated ES cell genotypes. *Nkx2-5*$^{lacZ/+}$↔wild-type chimaeras (D,G) were morphologically normal. *Nkx2-5*$^{GFP/lacZ}$↔wild-type (E,H) and Nkx2-5$^{GFP/Y\text{-}A}$↔wild-type (F,I) chimaeras showed arrested heart development and absence of ventricular trabeculation (arrows). (**J**,**K**) High-level chimaera at E9.5 carrying *Nkx2-5*$^{Y\text{-}A:IRESlacZ/+}$ ES cells (stained for lacZ) and section, showing relatively normal cardiac morphology. Abbreviations: CA, common atrium; LV, left ventricle; OT, outflow tract; PV, primitive ventricle; RA, right atrium.

Heterozygous Nkx2-5Y-A mutant chimaeras show congenital heart disease features

We were unable to derive a stable mouse line carrying one Y-A mutant allele using heterozygous Y-A ES cells (*Nkx2-5*$^{Y\text{-}AIRESlacZ/+}$), even though high-level chimaeras at 9.5-10.5 dpc with apparently normal general and cardiac morphology could be readily established (Fig. 6J,K). Only low-level chimaeras (<10% contribution) were found at birth and these failed to transmit the mutant allele through the germline after multiple matings. We therefore examined the morphology of *Nkx2-5*$^{Y\text{-}A:IRESlacZ/+(\Delta H)}$↔wild-type chimaeras at progressively later gestational stages, comparing them with *Nkx2-5*$^{lacZ/+(\Delta H)}$↔wild-type (heterozygous null) chimaeras as controls. We found that all moderate to high-level Y-A chimaeras examined at a gross level (n=25) succumbed just prior to birth from multiple cardiac malformations, while control chimaeras (n=50) showed normal morphology and survived beyond birth. Three independently derived clonal Y-A mutant ES cell lines gave identical results.

As early as 14.5 dpc, high-level Y-A mutant chimaeric hearts were enlarged, had a rounded apex and showed dilated ventricles and grossly dilated right atria (Fig. 7A-D). On their external surface, we observed unusual finger-like protrusions that were positive for *lacZ*, indicating that they were composed of mutant cells (Fig. 7B,C,G,H). Overall, Y-A mutant chimaeras expressed lower levels of *Myl2* (encoding myosin light chain 2v), known to be dependent upon Nkx2-5 during development (Lyons et al., 1995) (Fig. 7E,F). *Myl2* downregulation was exaggerated in patches that probably correspond to areas populated predominantly by mutant cells (Fig. 7E,F); indeed, the finger-like projections expressed lower levels of *Myl2* compared to surrounding myocardium (Fig. 7J, brackets). A blistering within the epicardium and between epicardium and myocardium was also a common feature (Fig. 7I, brackets). Histological examination showed that the right ventricular wall in Y-A chimaeras was thin and hypertrabeculated, reminiscent of ventricular non-compaction in humans (Pashmforoush et al., 2004) (Fig. 7K,L,M). The inter-ventricular septum was often fenestrated (Fig. 7L), or showed frank ventricular septal defect (VSD; n=5/8) (Fig. 7F,M). Transmural lesions of ventricular free wall were also occasionally detected (usually plugged with a blood clot) and these were surrounded by mutant cells (Fig. 7N; data not shown). Tricuspid valves were absent or hypoplastic (n=8/8) and, when present, were conical in shape with leaflets sitting below the annulus and partially fused to the interventricular septum or ventricular free wall (Fig. 7O,P). These are classical features of Ebstein's anomaly, which occurs occasionally in humans with *NKX2.5* mutations (Benson et al., 1999). Mitral valves were relatively normal. In the atria, the trabecular pattern was finer and more florid than in control hearts (Fig. 7E,F,H; data not shown), although this is likely to be secondary to dilation, as atrial trabeculae were not always composed of mutant cells (Fig. 7D).

Nkx2-5 has been suggested to confer resistance to cell death-inducing treatments (Monzen et al., 2002). To examine whether cell death due to compromised myocardium was a feature of the Y-A

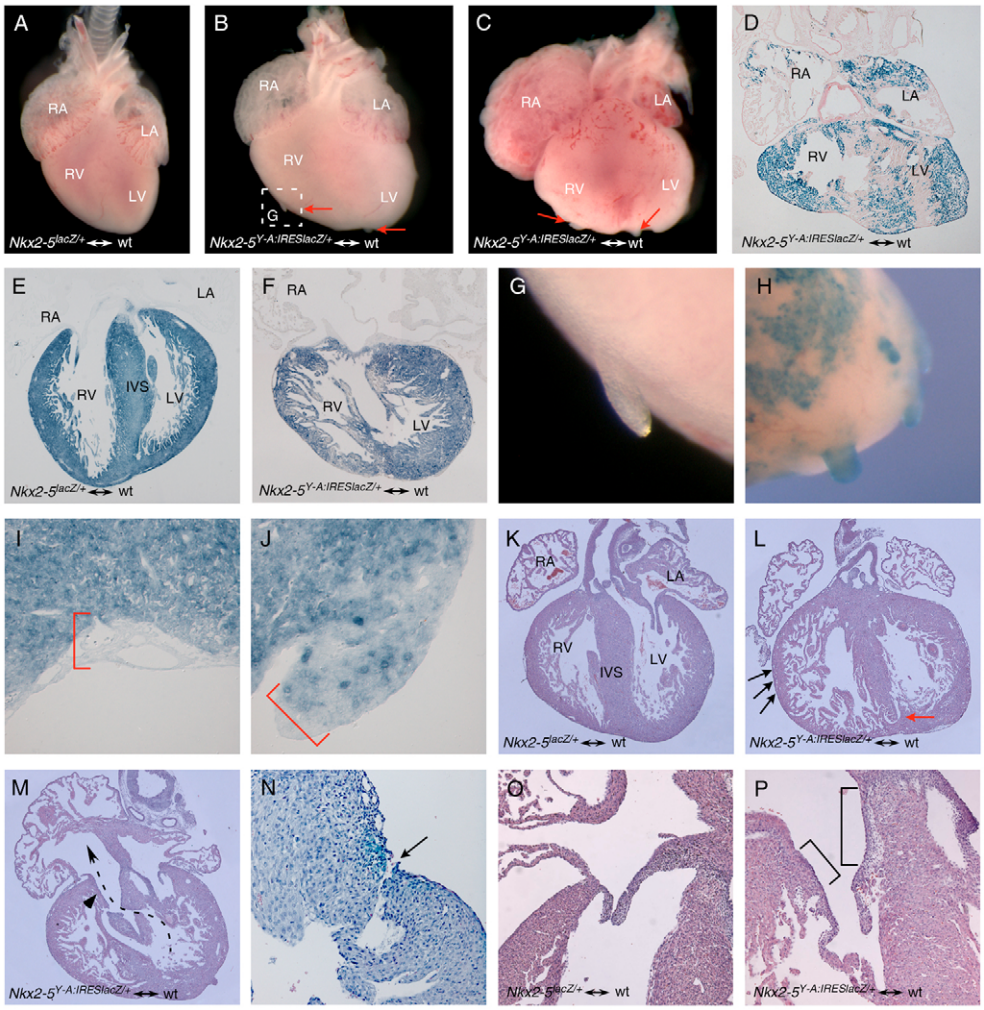

Fig. 7. Heterozygous Nkx2-5Y-A ES cell chimaeras show cardiac defects. (**A**) Control chimaeric heart at 14.5 dpc produced from an *Nkx2-5*$^{lacZ/+}$↔wild-type ES cells. (**B-D**) Three examples of Nkx2-5$^{GFP/Y-A}$↔wild-type chimaeras at 14.5 dpc, the section in D stained for *lacZ*. Arrows in B and C show finger-like projections from the myocardium. (**E**,**F**,**I**,**J**) Section in situ hybridisation for *Myl2* transcript showing low Myl2 transcript levels across the chimaeric heart (F) and in finger-like projections (J), and epicardial blistering (I). (**G**,**H**) Enlargements of boxed area in B showing finger-like projections that contain mutant Y-A cells (*lacZ* staining in H). (**K-P**) Histological examination of control *Nkx2-5*$^{lacZ/+}$↔wild-type chimaera (K,O) and mutant Nkx2-5$^{Y-A:IRESlacZ/+}$↔wild-type chimaeric hearts (L-N,P) highlighting right ventricular non-compaction (black arrows, L), fenestration of the ventricular septum (red arrow, L), ventricular septal defect (broken arrow in M), tricuspid valve showing Ebstein's anomaly (arrowhead, M; brackets, P) and transmural lesion (arrow, N). Abbreviations: IVS, interventricular septum; LA, left atrium; LV, left ventricle; RA, right atrium; RV, right ventricle.

mutant phenotype, we compared the prevalence of apoptosis in atria and ventricular free walls of *Nkx2-5*$^{Y-A:IRESlacZ/+(\Delta H)}$↔wild-type (Y-A heterozygous) and control *Nkx2-5*$^{lacZ/+(\Delta H)}$↔wild-type (null heterozygous) chimaeras at 14.5 dpc using the TUNEL assay. We chose this stage of analysis to be well before the systemic demise seen in mutant chimaeras. In the atria of Y-A chimaeras, TUNEL-positive cells were 3.34-fold more evident than in controls [4.93±0.59 (mean±s.e.m.) versus 1.48±0.22, P<0.005], while in the ventricular free walls there was a 3.24-fold difference (9.40±0.95 versus 2.90±0.27, P<0.008).

DISCUSSION

Nkx2-5 sits high in the cardiac regulatory hierarchy and can act as a transcriptional activator and repressor. As for its *Drosophila* cognate, Tinman, Nkx2-5 probably functions as an organ field-specific 'selector protein', defining a cardiac context for the multiple signal-dependent transcriptional processes acting in heart development (Halfon et al., 2000; Mann and Carroll, 2002). In this paper, we have defined a novel and essential tyrosine-rich domain within Nkx2-5 and describe a new model of Nkx2-5 deficiency that displays many phenotypic features of *NKX2.5* CHD in humans.

The YRD of mouse Nkx2-5 minimally encompasses a region of 37 amino acids with 24% (9/37) being tyrosine. A similar density of tyrosines, with occasional conservative substitutions to phenylalanine, are present in all known vertebrate and chordate members of the cardiac clade of NK2 homeodomain proteins. Other amino acids within the YRD are also highly conserved. However, the YRD and other conserved domains in the C terminus of Nkx2-5 (Fig. 2A), are lacking in *Drosophila* Tinman. We show here that in cardiac NK2 proteins from cephalopod molluscs, highly active invertebrates that possess a sophisticated closed cardiovascular system, the YRD and three of the conserved Nkx2-5 C-terminal signature domains are present. Thus, the YRD has co-evolved with the homeodomain since before divergence of the vertebrate and invertebrate lines some 550 million years ago. Tyrosine-rich regions, possibly ancestral YRDs, are recognisable C-terminal to the homeodomain in NK2-class proteins from *C. elegans*, mosquito, honeybee, hydra and sea anemone, the last two being members of the diploblastic phylum cnidaria (see Fig. 2A). A YRD-like element is also seen in *Drosophila* bagpipe-related genes. The YRD may therefore have a truly ancient origin in metazoan evolution. We have previously suggested that Tinman and Nkx2-5 have diverged mechanistically, as evidenced by the inability of Nkx2-5 to rescue heart development in Tinman mutant flies (Ranganayakulu et al., 1998). Our data now show that the domain structure of Tinman is in fact highly derived, highlighting a dramatic example of protein evolution within an otherwise conserved pathway. Protein evolution of this sort has been described in other selector gene pathways, for example, those involving *Drosophila* homeotic proteins Ubx and Hox3, and

segmentation protein fushi tarazu (Mann and Carroll, 2002). In Tinman, loss of C-terminal domains may have coincided with gain of an alternative cardiogenic mechanism involving an N-terminal domain (Ranganayakulu et al., 1998).

In mouse Nkx2-5, the YRD is essential for the earliest stages of cardiac development. Cardiomyocytes purified by FACS from EBs expressing only the Nkx2-5Y-A allele possessed the same gene expression signature defects as fully null cardiocytes. Importantly, chimaeras bearing a high investment of Y-A mutant cells showed a phenocopy of the genetic null mutant. The strongly compromised nature of the Y-A allele was also demonstrated by the phenotype of chimaeras carrying heterozygous *Nkx2-5*$^{Y\text{-}A:IRESlacZ/+}$ ES cells. These chimaeras died in the immediate pre-natal period with a constellation of severe heart malformations, whereas high-level chimaeras made from heterozygous *Nkx2-5*$^{lacZ/+}$ ES cells showed normal heart development. It is evident from these experiments that the Y-A mutation is effectively null in vivo, and in the presence of wild-type protein has a strong dominant-negative-like activity, producing phenotypes far stronger than haploinsufficiency. The potency of this effect is truly remarkable – while rare genes display haploinsufficiency in adults sufficient to cause difficulties in maintaining a genetically modified line (Bruneau et al., 2001), chimaeras carrying heterozygous Nkx2-5Y-A ES cells all succumbed at foetal stages.

It is noteworthy that in null and Y-A mutant EBs, there was an increase in the mean level of GFP fluorescence on FACS profiles relative to that seen in heterozygous null cells (Fig. 5B). Furthermore, in *Nkx2-5*$^{GFP/Y\text{-}A}$ mutant chimaeras, *lacZ* appeared to be upregulated dorsal to the heart compared with *Nkx2-5*$^{GFP/LacZ}$ chimaeras (Fig. 6E,F). Our recent microarray data on RNA from *Nkx2-5* heterozygous and null mutant hearts suggest that this increase is due to the loss of an Nkx2-5 negative auto-regulatory feedback loop in the nulls (O. Prall and R.P.H., unpublished).

One tyrosine codon within the YRD of human *NKX2.5* has been found to be mutated to a nonsense codon in a CHD family, with individuals over three generations displaying a spectrum of cardiac malformations identical to that seen in families with homeodomain mutations (Benson et al., 1999). This change disrupts the YRD and eliminates downstream amino acids. The characteristics of this mutation further highlight the importance of C-terminal domains for NKX2-5 function in vivo.

Although the YRD is essential for Nkx2-5 function in vivo, our experiments do not address directly the mechanism of its action. Nkx2-5 acts in positive synergy with other transcription factors to activate or repress expression of genes encoding both cardiac transcription factors and chamber differentiation proteins. These functions are consistent with a role for Nkx2-5 as a high-level selector protein (Mann and Carroll, 2002). The repressive activities of Nkx2-5 appear to be mediated through the TN-domain and C terminus. The TN domain has been shown to bind to the Groucho family of co-repressors (Muhr et al., 2001), whereas the C terminus carries an activity that represses the strong N-terminal trans-activation domain (Sepulveda et al., 1998). Owing to the presence of negative modulatory domains, the transcriptional activity of Nkx2-5 in vitro is therefore latent, but can be unmasked if all C-terminal amino acids are deleted, or when other cardiac transcription factors that directly interact with Nkx2-5 are co-expressed. Mutagenesis shows that the NK2SD acts as a negative modulator of Nkx2-5 transcription (Ranganayakulu et al., 1998; Sepulveda et al., 1998) and, when isolated from Nkx2-5, it can repress the strong trans-activation domains of the viral protein, VP16, in cis (Watada et al., 2000).

The YRD, which has positive transcriptional activity in vitro when fused to Gal4 DBD, may contribute to the positive activities of Nkx2-5. It could also serve to integrate its positive and negative functions. We note that the ability of Nkx2-5 C-terminal amino acids to repress the N-terminal trans-activation domain was intact in the YRD mutant, as was its ability to collaborate with GATA4/5, Tbx2/5/20, SRF and myocardin on the *Nppa* promoter (see Fig. S1 in the supplementary material). The activity of the Nkx2-5Y-A mutant on the *Gja5* promoter in the presence of Tbx20 and Gata5 was compromised, albeit only weakly (~50%). Therefore, the impact of YRD mutation within the context of Nkx2-5 in vitro is subtle at best. New assays that more robustly model its in vivo function await development.

Other homeoproteins possess accessory domains that are essential for their function and target specificity, and such inter-dependent relationships may have been necessary for expansion of the homeobox gene family and diversification of its functional repertoire. In the case of Pax transcription factors, for example, the paired domain and paired homeodomain act interdependently as a bipartite structure, with the paired domain influencing the DNA-binding properties of the homeodomain (Mann and Carroll, 2002). The YRD could also act in this way for NK2-class homeodomains, or attract accessory proteins that modulate chromatin or transcriptional activity in other ways.

Haploinsufficiency for *Nkx2-5* in mice leads to cardiac phenotypes that are much less severe than those seen in humans heterozygous for *NKX2.5* mutations: frank ASD occurs in 1% of mutant mice but in more than 70% of humans; conduction system defects, although evident in mice, are mild and do not progress to second and third degree conduction block as in humans (Biben et al., 2000). The significantly more severe human phenotypes may be due to subtle differences in the cardiac genetic program between species or the dominant-negative nature of some human mutations (Kasahara and Benson, 2004). Importantly, however, an animal model that accurately reflects the developmental defects seen in individuals with *NKX2.5* mutations is lacking. We show here that in chimaeric mice, the Nkx2-5Y-A allele produces a spectrum of cardiac phenotypes that overlaps with the more severe anomalies seen in the human disease. ASD, VSD, valvular anomalies, dilated cardiomyopathy and ventricular non-compaction are all components of *NKX2.5* mutant pathology. Increased myocyte death, presumably in response to reduced resistance to biomechanical stress, is also evident in chimaeras, and could be an unidentified component of the human disease. Indeed, we have recently also discovered a novel *NKX2.5* mutation in a human family with dilated cardiomyopathy, and ASD and conduction defects (Diane Fatkin and R.P.H., unpublished). Although molecular pathways remain to be elucidated, it is likely that the structural integrity of mutant myocytes and cellular adhesion between them is compromised in Nkx2-5Y-A chimaeras, as suggested by phenotypes involving downregulation of myofilament gene *Myl2*, transmural lesions, blistering between epicardium and myocardium, and the possible ‘streaming’ of myocytes into finger-like protrusions. We have recently developed a hypomorphic genetic model of Nkx2-5 function in mouse that has phenotypic manifestations similar to those seen in hearts of Y-A mutant chimaeras (O. Prall and R.P.H., unpublished). Our new models should offer valuable opportunities for further understanding of the complex and varied phenotypic manifestations of *NKX2.5* mutations. Furthermore, experiments focused on the specific role of the YRD should expand our understanding of cardiac developmental pathways and CHD.

We thank Mark Norman and Karen Crawford for cephalopod embryos, Owen Prall for help with PCR and FACS, Dee Lynch for help with ES cell culture, Stuart Gilchrist for statistical analysis, and Edouard Stanley for discussions. S.L.D. is a Pfizer Foundation of Australia Fellow and M.J.S. was a National Heart, Lung, and Blood Institute (USA) fellow.

Supplementary material

Supplementary material for this article is available at http://dev.biologists.org/cgi/content/full/133/7/1311/DC1

References

Azpiazu, N. and Frasch, M. (1993). *tinman* and *bagpipe*: two homeo box genes that determine cell fates in the dorsal mesoderm of *Drosophila*. *Genes Dev.* **7**, 1325-1340.

Bader, A., Gruss, A., Hollrigl, A., Al-Dubai, H., Capetanaki, Y. and Weitzer, G. (2001). Paracrine promotion of cardiomyogenesis in embryoid bodies by LIF modulated endoderm. *Differentiation* **68**, 31-43.

Baugh, L. R., Hill, A. A., Brown, E. L. and Hunter, C. P. (2001). Quantitative analysis of mRNA amplification by in vitro transcription. *Nucleic Acids Res.* **29**, E29.

Benson, D. W., Silberbach, G. M., Kavanaugh-McHugh, A., Cottrill, C., Zhang, Y., Riggs, S., Smalls, O., Johnson, M. C., Watson, M. S., Seidman, J. G. et al. (1999). Mutations in the cardiac transcription factor NKX2.5 affect diverse cardiac developmental pathways. *J. Clin. Invest.* **104**, 1567-1573.

Biben, C. and Harvey, R. P. (1997). Homeodomain factor Nkx2-5 controls left-right asymmetric expression of bHLH *eHand* during murine heart development. *Genes Dev.* **11**, 1357-1369.

Biben, C., Weber, R., Kesteven, S., Stanley, E., McDonald, L., Elliott, D. A., Barnett, L., Koentgen, F., Robb, L., Feneley, M. et al. (2000). Cardiac septal and valvular dysmorphogenesis in mice heterozygous for mutations in the homeobox gene Nkx2-5. *Circ. Res.* **87**, 888-895.

Bodmer, R. (1993). The gene *tinman* is required for specification of the heart and visceral muscles in *Drosophila*. *Development* **118**, 719-729.

Bruneau, B. G., Bao, Z.-Z., Tanaka, M., Schott, J.-J., Izumo, S., Cepko, C. L., Seidman, J. G. and Seidman, C. E. (2000). Cardiac expression of the ventricle-specific homeobox gene *Irx4* is modulated by Nkx2-5 and dHand. *Dev. Biol.* **217**, 266-277.

Bruneau, B. G., Nemer, G., Schmitt, J. P., Charron, F., Robitaille, L., Caron, S., Conner, D. A., Gessler, M., Nemer, M., Seidman, C. E. et al. (2001). A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. *Cell* **106**, 709-721.

Charron, F., Tsimiklis, G., Arcand, M., Robitaille, L., Liang, Q., Molkentin, J. D., Meloche, S. and Nemer, M. (2001). Tissue-specific GATA factors are transcriptional effectors of the small GTPase RhoA. *Genes Dev.* **15**, 2702-2719.

Chen, C. Y. and Schwartz, R. J. (1996). Recruitment of the tinman homolog Nkx-2.5 by serum response factor activates cardiac alpha-actin gene transcription. *Mol. Cell. Biol.* **16**, 6372-6384.

Chen, F., Kook, H., Milewski, R., Gitler, J. P., Lu, M. M., Li, J., Nazarian, R., Schnepp, R., Jen, K., Biben, C. et al. (2002). *Hop* is an unusual homeobox gene that modulates cardiac development. *Cell* **110**, 713-723.

Choi, C. Y., Lee, Y. M., Kim, Y. H., Park, T., Jeon, B. H., Schulz, R. A. and Kim, Y. (1999). The homeodomain transcription factor NK-4 acts as either a transcriptional activator or repressor and interacts with the p300 coactivator and the Groucho corepressor. *J. Biol. Chem.* **274**, 31543-31552.

Cripps, R. M. and Olson, E. N. (2002). Control of cardiac development by an evolutionarily conserved transcriptional network. *Dev. Biol.* **246**, 14-28.

Durocher, D., Charron, F., Warren, R., Schwartz, R. J. and Nemer, M. (1997). The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. *EMBO J.* **16**, 5687-5696.

Elliott, D. A., Kirk, E. P., Yeoh, T., Chandar, S., McKenzie, F., Taylor, P., Grossfeld, P., Fatkin, D., Jones, O., Hayes, P. et al. (2003). Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. *J. Am. Coll. Cardiol.* **41**, 2072-2076.

Engelmann, G. L. (1993). Coordinate gene expression during neonatal rat heart development. A possible role for the myocyte in extracellular matrix biogenesis and capillary angiogenesis. *Cardiovasc. Res.* **27**, 1598-1605.

Evans, S. (1999). Vertebrate tinman homologues and cardiac differentiation. *Semin. Cell Dev. Biol.* **10**, 73-83.

Fijnvandraat, A. C., van Ginneken, A. C., Schumacher, C. A., Boheler, K. R., Lekanne Deprez, R. H., Christoffels, V. M. and Moorman, A. F. (2003). Cardiomyocytes purified from differentiated embryonic stem cells exhibit characteristics of early chamber myocardium. *J. Mol. Cell. Cardiol.* **35**, 1461-1472.

Fye, W. B. (1987). The origin of the heart beat: a tale of frogs, jellyfish, and turtles. *Circulation* **76**, 493-500.

Grow, M. W. and Krieg, P. A. (1998). Tinman function is essential for vertebrate heart development: elimination of cardiac differentiation by dominant inhibitory mutants of the *tinman*-related genes, *XNkx2-3* and *XNkx2-5*. *Dev. Biol.* **204**, 187-196.

Habets, P. E., Moorman, A. F., Clout, D. E., van Roon, M. A., Lingbeek, M., van Lohuizen, M., Campione, M. and Christoffels, V. M. (2002). Cooperative action of Tbx2 and Nkx2.5 inhibits ANF expression in the atrioventricular canal: implications for cardiac chamber formation. *Genes Dev.* **16**, 1234-1246.

Halfon, M. S., Carmena, A., Gisselbrecht, S., Sackerson, C. M., Jimenez, F., Baylies, M. K. and Michaelson, A. M. (2000). Ras pathway specificity is determined by the integration of multiple signal-acivated and tissue-restricted transcription factors. *Cell* **103**, 63-74.

Harvey, R. P. (1996). NK-2 homeobox genes and heart development. *Dev. Biol.* **178**, 203-216.

Haun, C., Alexander, J., Stainier, D. Y. and Okkema, P. G. (1998). Rescue of caenorhabditis elegans pharyngeal development by a vertebrate heart specification gene. *Proc. Natl. Acad. Sci. USA* **95**, 5072-5075.

Hiroi, Y., Kudoh, S., Monzen, K., Ikeda, Y., Yazaki, Y., Nagai, R. and Komuro, I. (2001). Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. *Nat. Genet.* **28**, 276-280.

Holland, N. D., Venkatesh, T. V., Holland, L. Z., Jacobs, D. K. and Bodmer, R. (2003). AmphiNK2-tin, an amphioxus homeobox gene expressed in myocardial progenitors: insights into evolution of the vertebrate heart. *Dev. Biol.* **255**, 128-137.

Jay, P. Y., Harris, B. S., Maguire, C. T., Buerger, A., Wakimoto, H., Tanaka, M., Kuperschmidt, S., Roden, D. M., Schultheiss, T. M., O'Brien, T. X. et al. (2004). *Nkx2-5* mutation causes anatomic hypoplasia of the cardiac conduction system. *J. Clin. Invest.* **113**, 1130-1137.

Kasahara, H. and Benson, D. W. (2004). Biochemical analyses of eight *NKX2.5* homeodomain missense mutations causing atrioventricular block and cardiac anomalies. *Cardiovasc. Res.* **64**, 40-51.

Kasahara, H., Lee, B., Schott, J. J., Benson, D. W., Seidman, J. G., Seidman, C. E. and Izumo, S. (2000). Loss of function and inhibitory effects of human CSX/NKX2.5 homeoprotein mutations associated with congenital heart disease. *J. Clin. Invest.* **106**, 299-308.

Kortschak, R. D., Samuel, G., Saint, R. and Miller, D. J. (2003). EST analysis of the cnidarian Acropora millepora reveals extensive gene loss and rapid sequence divergence in the model invertebrates. *Curr. Biol.* **13**, 2190-2195.

Lyons, I., Parsons, L. M., Hartley, L., Li, R., Andrews, J. E., Robb, L. and Harvey, R. P. (1995). Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeobox gene *Nkx2-5*. *Genes Dev.* **9**, 1654-1666.

Mann, R. and Carroll, S. B. (2002). Molecular mechanisms of selector gene function and evolution. *Curr. Opin. Genet. Dev.* **12**, 592-600.

McElhinney, D. B., Geiger, E., Blinder, J., Benson, D. W. and Goldmuntz, E. (2003). NKX2.5 mutations in patients with congenital heart disease. *J. Am. Coll. Cardiol.* **42**, 1650-1655.

McKinsey, T. A., Zhang, C.-L., Lu, J. and Olson, E. N. (2000). Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. *Nature* **408**, 106-111.

Molkentin, J. D., Antos, C., Mercer, B., Taigen, T., Miano, J. M. and Olson, E. N. (2000). Direct activation of a *GATA6* cardiac enhancer by Nkx2.5: evidence for a reinforcing regulatory network of Nkx2.5 and GATA transcription factors in the developing heart. *Dev. Biol.* **217**, 301-309.

Monzen, K., Zhu, W., Kasai, H., Hiroi, Y., Hosoda, T., Akazawa, H., Zou, Y., Hayashi, D., Yamazaki, T., Nagai, R. et al. (2002). Dual effects of the homeobox transcription factor Csx/Nkx2-5 on cardiomyocytes. *Biochem. Biophys. Res. Commun.* **298**, 493-500.

Muhr, J., Andersson, E., Persson, M., Jessell, T. M. and Ericson, J. (2001). Groucho-mediated transcriptional repression establishes progenitor cell pattern and neuronal fate in the ventral neural tube. *Cell* **104**, 861-873.

Okkema, P. G., Ha, E., Haun, C., Chen, W. and Fire, A. (1997). The *Caenorhabditis elegans* NK-2 homeobox gene *ceh-22* activates pharyngeal muscle gene expression in combination with *pha-1* and is required for normal pharyngeal development. *Development* **124**, 3965-3973.

Park, M., Lewis, C., Turbay, D., Chung, A., Chen, J. N., Evans, S., Breitbart, R. E., Fishman, M. C., Izumo, S. and Bodmer, R. (1998). Differential rescue of visceral and cardiac defects in Drosophila by vertebrate tinman-related genes. *Proc. Natl. Acad. Sci. USA* **95**, 9366-9371.

Pashmforoush, M., Lu, J. T., Chen, H., St Amand, T., Kondo, R., Pradervand, S., Evans, S., Clark, B., Feramisco, J. R., Giles, W. et al. (2004). Nkx2-5 pathways and congenital heart disease: loss of ventricular myocyte lineage specification leads to progressive cardiomyopathy and complete heart block. *Cell* **117**, 373-386.

Raffin, M., Leong, L. M., Rones, M. S., Sparrow, D., Mohun, T. and Mercola, M. (2000). Subdivision of the cardiac Nkx2.5 expression domain into myogenic and nonmyogenic compartments. *Dev. Biol.* **218**, 326-340.

Ranganayakulu, G., Elliott, D. A., Harvey, R. P. and Olson, E. N. (1998). Divergent roles for NK-2 class homeobox genes in cardiogenesis in flies and mice. *Development* **125**, 3037-3048.

Sadowski, I., Bell, B., Broad, P. and Hollis, M. (1992). GAL4 fusion vectors for expression in yeast or mammalian cells. *Gene* **118**, 137-141.

Sepulveda, J. L., Belaguli, N., Nigam, V., Chen, C. Y., Nemer, M. and Schwartz, R. J. (1998). GATA-4 and Nkx-2.5 coactivate Nkx-2 DNA binding targets: role for regulating early cardiac gene expression. *Mol. Cell. Biol.* **18**, 3405-3415.

Shimizu, H. and Fujisawa, T. (2003). Peduncle of Hydra and the heart of higher organisms share a common ancestral origin. *Genesis* **36**, 182-186.

Shiratori, H., Sakuma, R., Watanabe, M., Hashiguchi, H., Mochida, K., Sakai, Y., Nishino, Y., Saijoh, Y., Whitman, M. and Hamada, H. (2001). Two-step regulation of left-right asymmetric expression of *Pitx2*: initiation by Nodal signalling and maintenance by Nkx2. *Mol. Cell* **7**, 139-149.

Stanley, E. G., Biben, C., Elefanty, A., Barnett, L., Koentgen, F., Robb, L. and Harvey, R. P. (2002). Efficient Cre-mediated deletion in cardiac progenitor cells conferred by a 3′UTR-ires-Cre allele of the homeobox gene Nkx2-5. *Int. J. Dev. Biol.* **46**, 431-439.

Stennard, F. A. and Harvey, R. P. (2005). T-box transcription factors and their roles in regulatory hierarchies in the developing heart. *Development* **132**, 4897-4910.

Stennard, F. A., Costa, M. W., Elliott, D. A., Rankin, S., Haast, S. J., Lai, D., McDonald, L. P., Niederreither, K., Dolle, P., Bruneau, B. G. et al. (2003). Cardiac T-box factor Tbx20 directly interacts with Nkx2-5, GATA4, and GATA5 in regulation of gene expression in the developing heart. *Dev. Biol.* **262**, 206-224.

Tsao, D. H., Gruschus, J. M., Wang, L.-H., Nirenberg, M. and Ferretti, J. A. (1994). Elongation of Helix III of the NK-2 Homeodomain upon Binding to DNA: A Secondary Structure Study by NMR. *Biochemistry (Mosc).* **33**, 15053-15060.

Ueyama, T., Kasahara, H., Ishiwata, T., Nie, Q. and Izumo, S. (2003). Myocardin expression is regulated by Nkx2.5, and its function is required for cardiomyogenesis. *Mol. Cell. Biol.* **23**, 9222-9232.

von Both, I., Silvestri, C., Erdemir, T., Lickert, H., Walls, J. R., Henkelman, R. M., Rossant, J., Harvey, R. P., Attisano, L. and Wrana, J. L. (2004). Foxh1 is essential for development of the anterior heart field. *Dev. Cell* **7**, 331-345.

Watada, H., Mirmira, R. G., Kalamaras, J. and German, M. S. (2000). Intramolecular control of transcriptional activity by the NK2-specific domain in NK-2 homeodomain proteins. *Proc. Natl. Acad. Sci. USA* **97**, 9443-9448.

Wells, M. J. (1992). The cephalopod heart: The evolution of a high-performance invertebrate pump. *Experientia* **48**, 800-808.

Yoshioka, H., Meno, C., Koshiba, K., Sugihara, M., Itoh, H., Ishimaru, Y., Inoue, T., Ohuchi, H., Semina, E. V., Murray, J. C. et al. (1998). *Ptx2*, a bicoid type homeobox gene, is involved in a lefty-signalling pathway in determination of left-right asymmetry. *Cell* **94**, 299-305.

Zou, Y., Evans, S., Chen, J., Kuo, H.-C., Harvey, R. P. and Chien, K. R. (1997). CARP, a cardiac ankyrin repeat protein, is downstream in the *Nkx2-5* homeobox gene pathway. *Development* **124**, 793-804.