

The RNA-binding specificity of the mouse Dazl protein

Julian P. Venables*, Matteo Ruggiu and Howard J. Cooke

Medical Research Council Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK

Received March 19, 2001; Revised and Accepted May 2, 2001

DBDJ/EMBL/GenBank accession nos*

ABSTRACT

DAZ is an RNA-binding protein encoded by a region on the Y chromosome implicated in infertility, and DAZ-like (Dazl) proteins are master regulators of germ line gene expression in all animals. In mice Dazl is only expressed in germ cells and is necessary for meiosis. A dual approach was taken to understand the RNA-binding specificity of the Dazl protein: (i) traditional SELEX and (ii) a novel tri-hybrid screen. Both approaches led to the same conclusion, namely that Dazl binds oligo(U) stretches interspersed by G or C residues. In a directed tri-hybrid assay the strongest interaction was with the consensus (GUn)_n. This motif is found in the 5' UTR of CDC25C whose homologue is thought to be the target of Boule, the Dazl homologue in flies. CDC25C 5' UTR also interacted specifically with Dazl *in vitro*. The tri-hybrid screen retrieved UTRs of known genes that may be physiological substrates of Dazl.

INTRODUCTION

The Dazl protein is ancestral to the human 'Deleted in Azoospermia' protein encoded by the Y chromosome Azoospermia region (AZFc) found deleted in some infertile men (1). The gene was transposed from an autosome after the divergence of the old-world monkeys and is solely autosomal in lower mammals (2). Dazl proteins are germ cell-specific and knocking them out across the animal kingdom has revealed various roles in germ cell differentiation. In *Xenopus* there is a primordial germ cell defect, which manifests as a failure to survive and migrate (3). There is male meiotic failure in *Drosophila* (4), female meiotic failure in *Caenorhabditis elegans* (5), and meiotic failure in both sexes of the Dazl knockout mouse (6).

Despite the many genetic studies aimed at elucidating the function of Dazl, its biochemical (i.e. RNA-binding) activity and the extent to which this is important for Dazl function have not been elucidated. We have used two independent screening methods, *in vitro* selection of high affinity RNA molecules bound by recombinant Dazl protein (SELEX) and selection of testes RNA sequences by tri-hybrid screening. Both

approaches yielded remarkably similar U-rich sequences. SELEX preferred short oligo(U) runs with intervening G residues, and the yeast tri-hybrid screen produced U-rich tracts with G or C residues in-between.

In *Drosophila* the Dazl (Boule) knockout has a block in G₂/M of meiosis 1 which can be overcome by expressing the reading frame from the meiosis-specific CDC25 homologue Twine. Expression of Twine correlates with the presence of Boule in the cytoplasm, leading the authors to propose all Dazl proteins as translational enhancers, although it is not established that Boule binds directly to either of Twine's UTRs (7). In mammals there are three CDC25 homologues that are differentially expressed; CDC25A is postulated to drive S phase, while CDC25B and CDC25C are involved in M phase of mitosis (8). Unlike the other CDC25 isoforms, CDC25C is predominantly expressed in germ cells, especially in spermatocytes and round spermatids in mice (9). Here we report that Dazl binds to the 5' UTR of CDC25C *in vitro* and in the yeast tri-hybrid system and that the (GUn)_n element in the 5' UTR of CDC25C that conforms to our consensus found by both SELEX and tri-hybrid screening is specifically bound by Dazl and therefore that CDC25C may be a target of Dazl in mice.

MATERIALS AND METHODS

GST fusion constructs

pGD encoding a GST fusion to mouse Dazl and pGDmut encoding a GST fusion of Dazl containing four point mutations in its RRM were created by cloning PCR products from pASd and pASmutRNP1/2 (10), respectively, into pGex 5X-1 (Pharmacia). PGD137 and pGD137mut encoding GST fusions of the first 137 amino acids (which includes the RRM) of wild-type and mutant Dazl were generated as above. Plasmids were transfected into *Escherichia coli* BL21 (DE3).

For SELEX, proteins were induced with 1 mM IPTG for 3 h. Cells were sonicated in PBS containing extra 250 mM KCl, 1 mM DTT and 0.05% Tween-20 and bound to glutathione-agarose. For the band shift assays, cells were induced with 0.2 mM IPTG and sonicated in denaturing extraction buffer (PBS containing additional 20 mM KCl, 1 mM DTT and 2% sarkosyl). The extract was bound to glutathione-agarose overnight at 4°C and beads were washed four times in extraction buffer with decreasing amounts of detergent to allow renaturation of the protein: (i) 2% sarkosyl; (ii) 0.75% sarkosyl, 0.75%

*To whom correspondence should be addressed at present address: Institute for Human Genetics, Ridley Building, Claremont Place, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK. Tel: +44 191 222 6827; Fax: +44 191 222 6662; Email: j.venables@ncl.ac.uk

Present address:

Matteo Ruggiu, Laboratory of Molecular Neuro-Oncology, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA

*AF347115, AJ409460–AJ409505

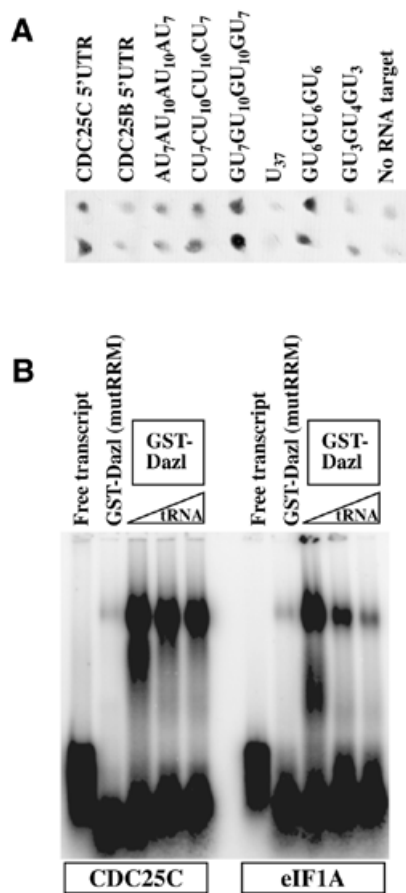


Figure 1. Binding of mouse Dazl to the GU-rich sequence in the 5' UTR of CDC25C. (A) Yeast tri-hybrid assays with two independent transformants from crosses between Dazl and the RNA sequences shown. (B) *In vitro* binding of the CDC25C 5' UTR to Dazl. Two 5' UTRs, CDC25C (left half) and eIF1A (right half) were transcribed *in vitro* and mixed with GST-Dazl. Lane 1 of each half shows free probe, and lane 2 of each half contains GST-Dazl with a mutant RRM and no tRNA. Lanes 3–5 of each half show complexes of GST-Dazl and RNA competed with 0–0.5 mg/ml tRNA.

paper until ~2 mM thick and dried in a gel drier at 60°C for 2 h, and subjected to phosphorimaging.

RESULTS

As finding RNA targets of known RNA-binding proteins has proven more challenging than identifying proteins that bind to known RNAs we have taken complimentary approaches to find the RNA targets of mouse Dazl in order to gain an insight into its mode of action.

In vitro selection of Dazl binding sites by SELEX

A GST fusion of the mouse Dazl protein was used to select a randomised pool of RNA 20mers for seven and nine rounds of SELEX. Twenty clones from the starting pool were sequenced to check for a random distribution of nucleotides. After seven and nine rounds of SELEX nearly identical nucleotide composition profiles were observed (54% U, 23/20% G, 12/14% C and 11/12% A) indicating that a consensus had been reached (Table S1). The sequences were all extremely U-rich (40–

75%) consistent with a previous homopolymer binding assay for mouse Dazl indicating a preference for poly(U) and poly(G) (14); indeed the only residue found flanked by 3 Us on either side was G. These data indicate a possible preference for GU-rich sequences.

Genetic screening for Dazl binding sites by yeast tri-hybrid assay

There are two yeast three-hybrid systems that have an RNA intermediate molecule that can be used for probing RNA–protein interactions. The more widely used relies on bacteriophage MS2 RNA binding to its coat protein (15) while the other uses tri-hybrid vector pREVRX that exploits the REV–RRE interaction (12,16,17). Neither system has been used to screen for unknown mRNA targets of RNA-binding proteins, but the idea was shown to work in theory when a U1 snRNP was used to retrieve U1 snRNA (18). However, there were several problems with the MS2 three-hybrid system, including high false positive rates and instability of longer inserts, so we decided to use the REV tri-hybrid system to screen for RNA targets of Dazl.

A mouse testes RNA library (average fragment size 250 bases) was constructed in pREVRX and screened using mouse Dazl as bait. Independent colonies (400 000) were screened for growth on 3-AT and subsequently for blue colour, yielding ~400 colonies that grew well on 3-AT selection, of which ~250 were blue in the β -Gal assay. Direct PCR and sequencing of positive colonies revealed >95% had oligo(U) sequences of >10 bases with interspersed G or C nucleotides (as with the SELEX a sample of sequences from the starting library showed no discernible bias). Plasmids from 52 yeast colonies were rescued, and when retested, 50 interacted with Dazl in the yeast tri-hybrid assay; when cotransfected with empty vector pACT, all were negative. Table S1 shows the 70 bases of sequence with the highest U content from the first 45 clones that gave a strong blue colour (full-length accession nos AJ409460–AJ409505). This sample is composed of 58% U residues and 13%, 15% and 14% A, C and G residues, respectively. However the nucleotide proportions found flanked by 3 Us in the sequence U₃NU₃ are 4% A, 51% C and 45% G, respectively, indicating a consensus minimal binding site of U₃G/CU₃. As the library was non-directionally cloned, Blast searching of the retrieved sequences revealed mostly matches with antisense sequences and ESTs. There were no matches to coding regions of any known genes but there were matches to four intronic sequences including double strand break protein spo11 (19) and mitotic protein Pin1 (20) and two 3' UTRs sequences: SNF-1 like kinase (containing GUUUUGUUUUGUUUUCGUUUUUUUUUUUUUUUUUUU) (21) and myristoylated alanine-rich C kinase substrate (containing UUUUUUGUUUUUUGUUUUUUUU) (22) and one 5' UTR, that of the tyrosine phosphatase PRL-1 (containing GUUGUUUUUUUGGUUUUUGCUUUUUUUUUUUUUUUUUUUUUUUUCGUUUUUGGUUUU) (23,24).

Dazl binds broken oligo(U) rich tracts with an order of preference G > C > A for the interspersing nucleotides

As the two methods of screening above gave similar results, the UTRs of the three mouse CDC25 homologues A, B and C were searched for broken oligo(U) tracts since the UTRs of CDC25 homologue 'Twine' in *Drosophila* are probable

targets for the Dazl homologue 'Boule' (7). The most striking sequence by far was a 37mer from the CDC25C 5' UTR GU₇GU₁₀GU₁₀GU₇. This sequence was not retrieved in the screen and was not in the library when checked by PCR. This may be because its 367 nt are theoretically able to form into a long stable hairpin structure (not shown) that may reduce efficiency of cloning.

The 5' UTRs of CDC25C, and as a control CDC25B, were cloned (by nested RT-PCR) into the tri-hybrid vector as was the 37mer sequence and variants of it in which the G residues were changed or the size and number of the repeats were reduced. A strong yeast tri-hybrid reaction was observed between Dazl and CDC25C 5' UTR but not with CDC25B 5' UTR (Fig. 1A). A strong interaction was also observed with the GU₇GU₁₀GU₁₀GU₇ sequence. When the Gs in this sequence were mutated to C the binding was slightly reduced. Only a weak interaction was observed when G was replaced by A and little if any reaction was observed when G was replaced by U, i.e. with an unbroken run of U residues. Binding of Dazl to reduced tract GU₆GU₆GU₆ was strong but binding to GU₃GU₄GU₃ was significantly weaker.

Dazl binds specifically to the 5' UTR of CDC25C *in vitro*

To confirm the tri-hybrid interactions between Dazl and the 5' UTR of CDC25C, radiolabelled 5' RACE products were transcribed *in vitro* and bound to GST-Dazl and run on an agarose gel. CDC25C 5' UTR was retarded by GST-Dazl in the presence of high concentrations of competitor tRNA (Fig. 1B) whereas the binding to the control 5' UTR eIF1A that does not contain oligo(U) sequences was significantly competed. GST-Dazl with four point mutations in its RRM failed to bind RNA. Together these results show that Dazl binds CDC25C 5' UTR specifically through its RRM domain.

DISCUSSION

Previous studies of the RNA-binding specificity of Dazl proteins showed binding to poly(U) and poly(G) but not poly(C) or poly(A) (14,25). This study further refines the RNA-binding specificity of Dazl to the consensus (G/CUn)_n and suggests specific targets that may be translationally regulated. Translational control is known to be an important part of the process of spermatogenesis (26), and intriguingly Dazl has been found associated with actively translating ribosomes (14). *Drosophila* Dazl (Boule) upregulates the translation of a meiosis-specific homologue of CDC25C (Twine) although it remains to be determined if this effect is direct (7). Support for the translational regulation of CDC25C mRNA comes from the observation that fresh translation of CDC25C is necessary for meiosis in pig oocytes (27) and *cdc25+* is translationally regulated by U-rich sequences in its 5' UTR in yeast (28). The fact that Dazl binds a U-rich element in the 5' UTR of CDC25C and that both genes are predominantly expressed in the male germline makes it attractive that Dazl may control the translation of CDC25C. However, while in *Drosophila* the control of CDC25 (Twine) translation is thought to be the primary function of Dazl (Boule) (7), in the Dazl knockout mouse the phenotype is detectable many days earlier (H.J.Cooke and P.Saunders, manuscript in preparation) than the CDC25C message and protein which are first detected around day 17 post partum, coincident with the first wave of

meiotic division (9). It is probable therefore that Dazl, which is present throughout germ cell development, has other earlier targets involved in commitment to enter meiosis.

A search through 7400 mouse 5' UTRs (29), revealed 34 possible downstream targets of Dazl (not shown) indicating that such sequences occur relatively frequently and therefore that Dazl may target multiple RNAs, but it remains to be established if any of them are transcribed in germ cells. Cell cycle regulator p27kip contains a (GUn)_n element in its 5' UTR that may be important for translational control (30) although in testes p27kip is predominantly expressed in Sertoli cells rather than germ cells (31). Although different antibodies have given opposing views about whether mouse Dazl is strictly cytoplasmic (32,33) in *Drosophila*, nuclear localisation is not essential for meiosis (34), so it is questionable whether the intronic sequences retrieved here are physiological sites of action.

If Dazl regulates translation it may do so alone or by interacting with the translation machinery through testes-specific proteins like the human 'DAZ associated protein' (35), which may be essential components of any assay systems. However, proof of the function of Dazl may await identification of its first essential targets in development. To this end we have shown the promise of the yeast tri-hybrid system to reveal mRNA sequences bound by RNA-binding proteins, and saturating screens of physiologically relevant substrates may be facilitated by refining the source of RNA for library construction.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

Thanks to Mary Taggart for technical assistance and Nicola Gray, Janet Partridge, Michael Ladomery and Nicholas Hastie for comments and advice. M.R. was supported by a Wellbeing fellowship.

REFERENCES

- Saxena,R., Brown,L.G., Hawkins,T., Alagappan,R.K., Skaletsky,H., Reeve,M.P., Reijo,R., Rozen,S., Dinulos,M.B., Distech,C.M. and Page,D.C. (1996) The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned. *Nat. Genet.*, **14**, 292-299.
- Cooke,H.J., Lee,M., Kerr,S. and Ruggiu,M. (1996) A murine homologue of the human DAZ gene is autosomal and expressed only in male and female gonads. *Hum. Mol. Genet.*, **5**, 513-516.
- Houston,D.W. and King,M.L. (2000) A critical role for Xdazl, a germ plasm-localized RNA, in the differentiation of primordial germ cells in *Xenopus*. *Development*, **127**, 447-456.
- Eberhart,C.G., Maines,J.Z. and Wasserman,S.A. (1996) Meiotic cell cycle requirement for a fly homologue of human Deleted in Azoospermia. *Nature*, **381**, 783-785.
- Karashima,T., Sugimoto,A. and Yamamoto,M. (2000) *Caenorhabditis elegans* homologue of the human azoospermia factor DAZ is required for oogenesis but not for spermatogenesis. *Development*, **127**, 1069-1079.
- Ruggiu,M., Speed,R., Taggart,M., McKay,S.J., Kilanowski,F., Saunders,P., Dorin,J. and Cooke,H.J. (1997) The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis. *Nature*, **389**, 73-77.
- Maines,J.Z. and Wasserman,S.A. (1999) Post-transcriptional regulation of the meiotic Cdc25 protein Twine by the Dazl orthologue Boule. *Nat. Cell Biol.*, **1**, 171-174.

8. Nilsson,I. and Hoffmann,I. (2000) Cell cycle regulation by the Cdc25 phosphatase family. *Prog. Cell Cycle Res.*, **4**, 107–114.
9. Wu,S. and Wolgemuth,D.J. (1995) The distinct and developmentally regulated patterns of expression of members of the mouse Cdc25 gene family suggest differential functions during gametogenesis. *Dev. Biol.*, **170**, 195–206.
10. Ruggiu,M. and Cooke,H.J. (2000) *In vivo* and *in vitro* analysis of homodimerisation activity of the mouse Dazl1 protein. *Gene*, **252**, 119–126.
11. Conrad,R.C., Giver,L., Tian,Y. and Ellington,A.D. (1996) *In vitro* selection of nucleic acid aptamers that bind proteins. *Methods Enzymol.*, **267**, 336–367.
12. Putz,U., Kremerskothen,J., Skehel,P. and Kuhl,D. (2000) RNA–protein interactions reconstituted by a tri-hybrid system. In Zhu,L. and Hannon,G.J. (eds), *Yeast Hybrid Technologies*. Eaton, Natick, MA, pp. 207–219.
13. Davis,W. and Schultz,R.M. (2000) Developmental change in TATA-box utilization during preimplantation mouse development. *Dev. Biol.*, **218**, 275–283.
14. Tsui,S., Dai,T., Warren,S.T., Salido,E.C. and Yen,P.H. (2000) Association of the mouse infertility factor DAZL1 with actively translating polyribosomes. *Biol. Reprod.*, **62**, 1655–1660.
15. Sengupta,D.J., Zhang,B., Kraemer,B., Pochart,P., Fields,S. and Wickens,M. (1996) A three-hybrid system to detect RNA–protein interactions *in vivo*. *Proc. Natl Acad. Sci. USA*, **93**, 8496–8501.
16. Putz,U., Skehel,P. and Kuhl,D. (1996) A tri-hybrid system for the analysis and detection of RNA–protein interactions. *Nucleic Acids Res.*, **24**, 4838–4840.
17. Monshausen,M., Putz,U., Rehbein,M., Schweizer,M., DesGroseillers,L., Kuhl,D., Richter,D. and Kindler,S. (2001) Two rat brain Staufein isoforms differentially bind RNA. *J. Neurochem.*, **76**, 155–165.
18. Sengupta,D.J., Wickens,M. and Fields,S. (1999) Identification of RNAs that bind to a specific protein using the yeast three-hybrid system. *RNA*, **5**, 596–601.
19. Romanienko,P.J. and Camerini-Otero,R.D. (2000) The mouse spo11 gene is required for meiotic chromosome synapsis. *Mol. Cell*, **6**, 975–987.
20. Fujimori,F., Takahashi,K., Uchida,C. and Uchida,T. (1999) Mice lacking Pin1 develop normally, but are defective in entering cell cycle from G(0) arrest. *Biochem. Biophys. Res. Commun.*, **265**, 658–663.
21. Ruiz,J.C., Conlon,F.L. and Robertson,E.J. (1994) Identification of novel protein kinases expressed in the myocardium of the developing mouse heart. *Mech. Dev.*, **48**, 153–164.
22. Seykora,J.T., Ravetch,J.V. and Aderem,A. (1991) Cloning and molecular characterization of the murine macrophage ‘68-kDa’ protein kinase C substrate and its regulation by bacterial lipopolysaccharide. *Proc. Natl Acad. Sci. USA*, **88**, 2505–2509.
23. Rundle,C.H. and Kappen,C. (1999) Developmental expression of the murine Prl-1 protein tyrosine phosphatase gene. *J. Exp. Zool.*, **283**, 612–617.
24. Diamond,R.H., Cressman,D.E., Laz,T.M., Abrams,C.S. and Taub,R. (1994) PRL-1, a unique nuclear protein tyrosine phosphatase, affects cell growth. *Mol. Cell. Biol.*, **14**, 3752–3762.
25. Houston,D.W., Zhang,J., Maines,J.Z., Wasserman,S.A. and King,M.L. (1998) A *Xenopus* DAZ-like gene encodes an RNA component of germ plasm and is a functional homologue of *Drosophila* boule. *Development*, **125**, 171–180.
26. Braun,R.E. (2000) Post-transcriptional control and male infertility. *Results Probl. Cell Differ.*, **28**, 111–129.
27. Dai,Y., Lee,C., Hutchings,A., Sun,Y. and Moor,R. (2000) Selective requirement for Cdc25C protein synthesis during meiotic progression in porcine oocytes. *Biol. Reprod.*, **62**, 519–532.
28. Daga,R.R. and Jimenez,J. (1999) Translational control of the cdc25 cell cycle phosphatase: a molecular mechanism coupling mitosis to cell growth. *J. Cell Sci.*, **112**, 3137–3146.
29. Pesole,G., Liuni,S., Grillo,G., Licciulli,F., Larizza,A., Makalowski,W. and Saccone,C. (2000) UTRdb and UTRsite: specialized databases of sequences and functional elements of 5' and 3' untranslated regions of eukaryotic mRNAs. *Nucleic Acids Res.*, **28**, 193–196.
30. Millard,S.S., Vidal,A., Markus,M. and Koff,A. (2000) A U-rich element in the 5' untranslated region is necessary for the translation of p27 mRNA. *Mol. Cell. Biol.*, **20**, 5947–5959.
31. Beumer,T.L., Kiyokawa,H., Roepers-Gajadien,H.L., van den Bos,L.A., Lock,T.M., Gademan,I.S., Rutgers,D.H., Koff,A. and de Rooij,D.G. (1999) Regulatory role of p27kip1 in the mouse and human testis. *Endocrinology*, **140**, 1834–1840.
32. Reijo,R.A., Dorfman,D.M., Slee,R., Renshaw,A.A., Loughlin,K.R., Cooke,H. and Page,D.C. (2000) DAZ family proteins exist throughout male germ cell development and transit from nucleus to cytoplasm at meiosis in humans and mice. *Biol. Reprod.*, **63**, 1490–1496.
33. Ruggiu,M., Saunders,P.T. and Cooke,H.J. (2000) Dynamic subcellular distribution of the DAZL protein is confined to primate male germ cells. *J. Androl.*, **21**, 470–477.
34. Cheng,M.H., Maines,J.Z. and Wasserman,S.A. (1998) Biphasic subcellular localization of the DAZL-related protein boule in *Drosophila* spermatogenesis. *Dev. Biol.*, **204**, 567–576.
35. Tsui,S., Dai,T., Roettger,S., Schempp,W., Salido,E.C. and Yen,P.H. (2000) Identification of two novel proteins that interact with germ-cell-specific RNA-binding proteins DAZ and DAZL1. *Genomics*, **65**, 266–273.