

# Dyskerin localizes to the nucleolus and its mislocalization is unlikely to play a role in the pathogenesis of dyskeratosis congenita

Nina S. Heiss, Andreas Girod<sup>1</sup>, Rüdiger Salowsky, Stefan Wiemann, Rainer Pepperkok<sup>1</sup> and Annemarie Poustka<sup>+</sup>

Deutsches Krebsforschungszentrum (DKFZ), Department of Molecular Genome Analysis, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany and <sup>1</sup>European Molecular Biology Laboratories (EMBL), Cell Biophysics, Meyerhofstrasse 1, D-69117 Heidelberg, Germany

Received August 3, 1999; Revised and Accepted September 21, 1999

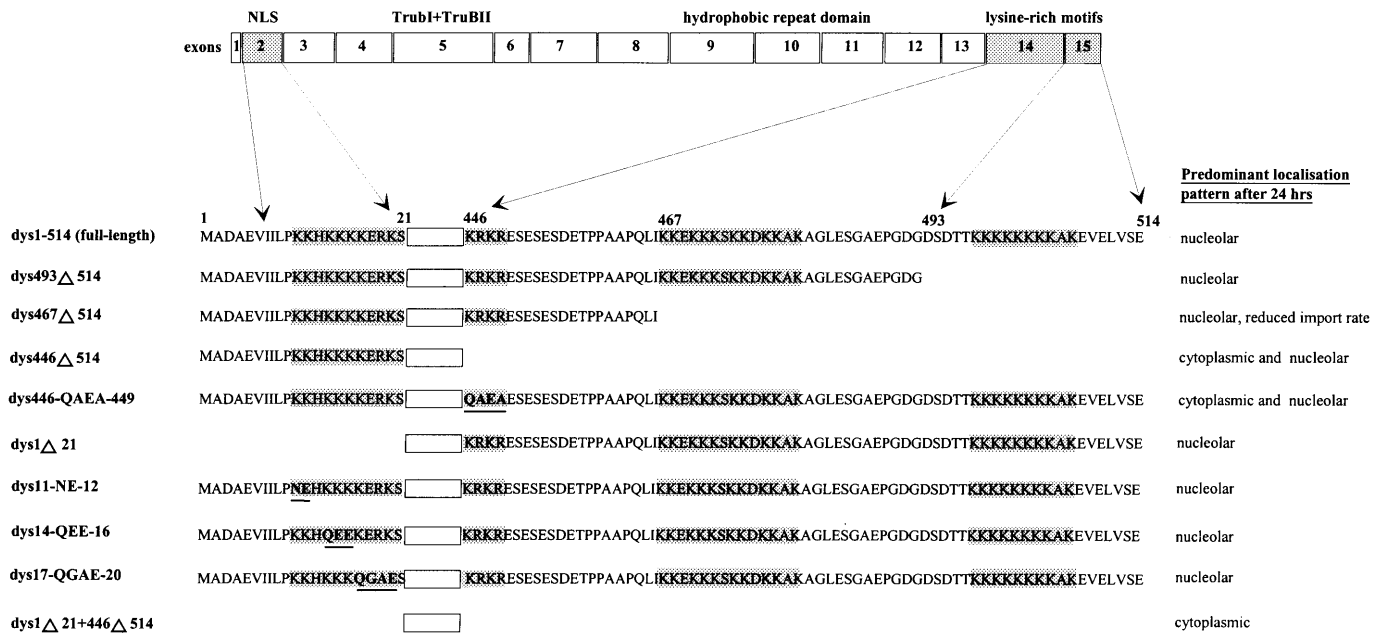
Mutations in the *DKC1* gene are responsible for causing the bone marrow failure syndrome, dyskeratosis congenita (DKC; OMIM 305000). The majority of mutations identified to date are missense mutations and are clustered in exons 3, 4 and 11. It is predicted that the corresponding protein dyskerin is a nucleolar phosphoprotein which functions in both pseudouridylation and cleavage of precursor rRNA. Dyskerin contains multiple putative nuclear localization signals (NLSs) at the N-terminus (KKHKKKKERKS) and C-terminus [KRKR(X)<sub>17</sub>KKEKKSKDKKAK(X)<sub>17</sub>KKKKKKKAKEVELVSE]. By fusing dyskerin with the enhanced green fluorescent protein (EGFP) and by following a time course of expression in mammalian cell lines, we showed that full-length dyskerin initially localizes to the nucleoplasm and subsequently accumulates in the nucleoli. A co-localization to the coiled bodies was observed in some cells where dyskerin–EGFP had translocated to the nucleoli. Analysis of a series of mutant constructs indicated that whereas the most C-terminal lysine-rich clusters [KKEKKSKDKKAK(X)<sub>17</sub>KKKKKKKAKEVELVSE] influence the rate of nucleoplasmic and nucleolar accumulation, the KRKR sequence is primarily responsible for the nuclear import. Nucleolar localization was maintained when either the N- or C-terminal motifs were mutated, but not when all NLSs were removed. We conclude that the intranuclear localization of dyskerin is accomplished by the synergistic effect of a number of NLSs and that the nucleolar localization signals are contained within the NLSs. Further, examination of dyskerin–EGFP fusions mimicking mutations detected in patients indicated that the intracellular mislocalization of dyskerin is unlikely to cause DKC.

## INTRODUCTION

X-linked recessive dyskeratosis congenita (DKC; OMIM 305000) is a congenital multisystem disorder characterized by nail dystrophy, mucosal leukoplakias and reticulate skin hyperpigmentation (1). The severity of the disease is marked by the concomitant development of pancytopenia, and progressive bone marrow failure leads to death within the first or second decade of life in 90% of cases (2). The patients also have an increased predisposition towards developing a range of malignancies. The gene responsible for causing the X-linked recessive form of the disease, *DKC1*, lies in distal Xq28 in very close proximity to the *MPP1* gene (3). The genomic sequence was determined and has facilitated an efficient screening for mutations across the 15 exons including the exon–intron boundaries. So far, 14 different missense mutations, a putative splice site mutation, a single amino acid deletion and a deletion of the last exon, exon 15, have been identified in DKC patients (3–5). As DKC affects mainly the rapidly dividing cells of the epithelia and the haemopoietic system, and only the cells expressing the normal allele survive in DKC carrier females (6,7), it is expected that the protein dyskerin functions in maintaining cell proliferation and/or survival.

Dyskerin reveals a high degree of evolutionary conservation and is the orthologue of the yeast Cbf5p, *Drosophila* Nop60B and rat NAP57 proteins (3,8–12). That dyskerin is a nucleolar protein and performs a role in rRNA biogenesis is supported by functional studies on the orthologous proteins. Precursor rRNA transcripts undergo a number of post-transcriptional modifications prior to packaging with ribosomal proteins. This includes the pseudouridylation of selected uridine residues, the site-specificity of which is guided by a class of small nucleolar RNAs (snoRNAs) called the box H + ACA snoRNAs (13). In yeast, the box H + ACA snoRNAs form a complex with a number of nucleolar proteins including Cbf5p, Gar1p, Nhp2p and Nop10p, where Cbf5p is the core component (14,15). Yeast Cbf5p and *Drosophila* Nop60B stabilize both the RNA and protein components of the box H + ACA complex. There is also indirect evidence that they represent a new member of the rRNA pseudouridine synthases (12,14). This agrees with the presence of two conserved TruB motifs in dyskerin which are believed to represent the uridine-binding motifs (16). The

<sup>+</sup>To whom correspondence should be addressed. Tel: +49 6221 42 3409; Fax: +49 6221 42 3454; Email: a.poustka@dkfz-heidelberg.de



**Figure 1.** Dyskerin-EGFP deletion and *in vitro* mutagenized constructs. The top bar represents full-length dyskerin: exons are numbered 1–15; exons containing the NLSs are shown in shaded boxes; TruBI + TruBII indicates the putative pseudouridine synthase domains. The amino acid sequences of the N- and C-terminal NLSs are shown below for clarity, the lysine-rich clusters are shaded, numbers refer to amino acid positions. For the deletion constructs, the EGFP tag was placed at either the N- or C-terminus. The amino acids modified by *in vitro* mutagenesis are shown in bold and underlined.

potential involvement of dyskerin in the cleavage of pre-rRNA derives from analyses on the Cbf5p and Nop60B mutants which accumulate the large precursor rRNAs at the expense of reduced levels of the cleaved 28S, 18S and 5.8S rRNA forms (12,14). The mature rRNA transcripts are packaged with ribosomal proteins and are then exported into the cytoplasm. The interacting protein of rat NAP57, Nopp140, shuttles between the cytoplasm and the nucleus, and a chaperoning role for NAP57 in the transport of ribosomal proteins was therefore proposed (8,17). It is conceivable that dyskerin may also play such a role.

The nucleus is a highly compartmentalized structure. In addition to the nucleoli, it contains a variety of other small subnuclear structures with a punctate distribution, such as the coiled bodies. The coiled bodies share a number of protein components which include p80coilin, Cbf5p/NAP57 and Nopp140 (18,19). Although the exact function of the coiled bodies is unknown, they are believed to play a role in the storage, maturation and transport of proteins necessary for rRNA processing, and are physically, dynamically and functionally linked with the nucleoli (18–21).

Many nuclear proteins are actively transported from the cytoplasm into the nucleus across the nuclear pore complex (NPC). This requires protein–protein interactions which are mediated by short clusters of sequences termed nuclear localization signals (NLSs). Most NLSs consist of basic amino acid residues, especially lysine and arginine, although the spatial distribution of acidic and neutral amino acids is also important (22–28). Dyskerin harbours a number of putative NLSs at the N- and C-termini. In particular, the C-terminal motif KRKR(X)<sub>17</sub>-KKEKKKSKDKKAK(X)<sub>17</sub>-KKKKKKKAKEVELVSE is reminiscent of overlapping bipartite NLS motifs that have been described for a number of nuclear proteins (22).

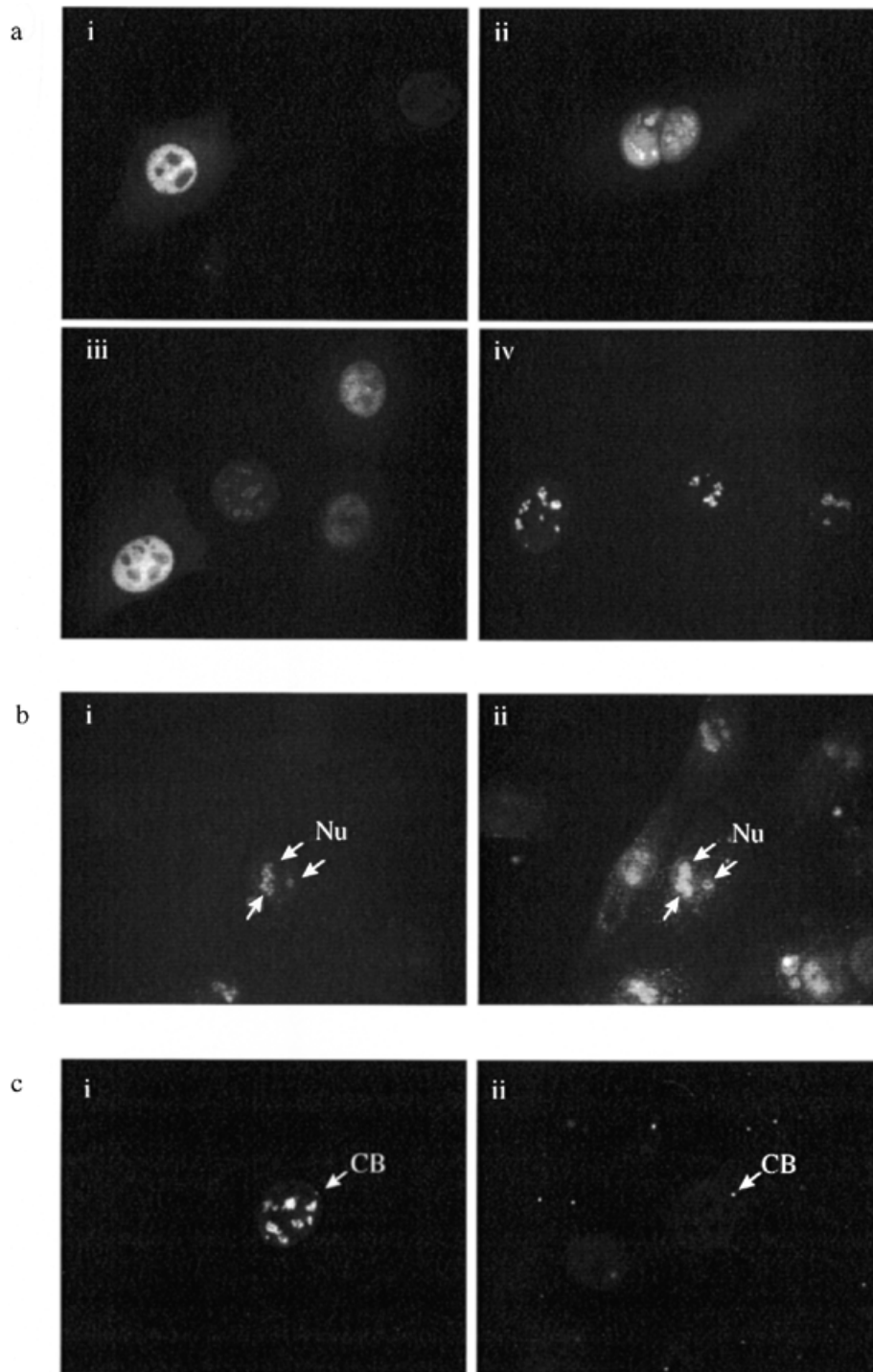
The functional investigations on yeast Cbf5p, *Drosophila* Nop60B and rat NAP57 viewed in conjunction with the predicted motifs in dyskerin all converge to support its nucleolar functions. However, to date, no functional analyses have been reported. To begin a functional characterization of dyskerin, we determined the intracellular localization by expressing dyskerin fused to the enhanced green fluorescent protein (EGFP) in mammalian cell lines. We also aimed functionally to delineate motifs which are responsible for the nuclear/nucleolar localization of dyskerin and to determine whether mislocalization could be one mechanism to explain the pathogenesis of DKC.

## RESULTS

### Dyskerin localizes to the nucleoli and the coiled bodies

To determine the intracellular localization of dyskerin, the full-length open reading frame was fused to EGFP at either the N- or C-terminus and transfected into Vero, NIH-3T3 and HeLa cells (dys1–514; Fig. 1). The expression was monitored at 24, 36, 48 and 72 h after transfection. At each time point, a heterogeneous localization pattern of dyskerin-EGFP was evident. Basically three types of localization were observed: (i) only nucleoplasmic; (ii) nucleoplasmic and nucleolar; and (iii) only nucleolar. There was an increase in the proportion of cells exhibiting only nucleolar localization with increasing time after transfection for both the N- and C-terminally tagged constructs. This indicated that dyskerin initially localizes to the nucleoplasm and then translocates to the nucleoli where it accumulates.

To demonstrate this effect over a more defined time course, the constructs were microinjected into the nuclei of Vero and NIH-3T3 cells. It was possible to microinject 200 cells per construct within 20 min. This provided a time interval of sufficient



**Figure 2.** (a) Representative cells showing the localization of full-length dyskerin (dys1–514) tagged to EGFP at the N- or C-terminus. (i) Nucleoplasmic localization observed in all cells 2 h after microinjection; (ii) nucleoplasmic and nucleolar localization observed in most cells 4, 6 and 8 h after microinjection, and in very few cells 24 h after microinjection; (iii) mixed population of cells observed 4, 6 and 8 h after microinjection, sometimes 24 h after microinjection; (iv) sole nucleolar localization observed in a few cells 4, 6 and 8 h after microinjection and representing the majority of cells after 24 h. (b) (i) Nucleolar localization of dys1–514; (ii) indirect immunofluorescence of the same cells stained with the B23 (nucleophosmin) antibody verifying the nucleoli. (c) (i) Localization of dys1–514 to the nucleoli and a coiled body; (ii) indirect immunofluorescence of the same cells stained with the p80coilin antibody verifying the position of the coiled body. Similar results were obtained for the construct mimicking the exon 15 patient deletion (dys493Δ514). Nu, nucleolus; CB, coiled body.

length for detailed documentation and imaging of at least 50 cells for each construct at exactly 2, 4, 6, 8 and 24 h after

microinjection. After 2 h, dyskerin localized only to the nucleoplasm (Fig. 2a, i) and then began to accumulate in the nucleolus

between 4 and 8 h after microinjection (Fig. 2a, ii and iii). Analogously to the transfections, the proportion of cells showing nucleolar localization increased over time. A complete translocation to the nucleoli occurred in the majority of cells after 24 h, and a mixed localization pattern was observed in only a few cells (Fig. 2a, iii and iv). This confirmed that the destination of dyskerin-EGFP is the nucleolus.

The nucleolar localization was verified by staining the cells with the B23 antibody specific for the nucleolar protein, nucleophosmin (Fig. 2b). In addition to the nucleolar localization, a few cells exhibited bright foci of fluorescence in the nucleoplasm. To identify these extranucleolar structures, cells expressing dyskerin-EGFP for 24 h were stained with the p80coilin antibody. This indicated that dyskerin co-localizes with the coiled bodies (Fig. 2c). An association of the orthologous rat NAP57 protein with the coiled bodies has already been reported (8). The localization of dyskerin is in agreement with the predicted functions.

### Mislocalization of dyskerin is unlikely to cause dyskeratosis congenita

Short clusters of basic residues are necessary, and in some cases sufficient, to direct a protein to the nucleus, although no strict consensus sequence exists (22). The C-terminus of dyskerin consists of lysine-rich clusters which contain a number of putative overlapping NLSs (Fig. 1). The exon 15 deletion identified in one DKC patient removes the C-terminal 22 amino acids, and this includes the most C-terminal lysine-rich cluster (Fig. 1; dys493 $\Delta$ 514) (3,5). We therefore sought to determine whether the sequence deleted in the patient is required for targeting dyskerin to the nucleus and/or the nucleoli and whether its removal causes dyskerin to mislocalize. Construct dys493 $\Delta$ 514 was transfected and microinjected into the various cell lines and its localization compared with that of the full-length construct at the same time intervals. No obvious spacial or temporal difference in the nucleoplasmic/nucleolar localization pattern was evident, and a co-localization to the coiled bodies was also observed (Fig. 2).

With a few exceptions, DKC is caused predominantly by missense mutations in the *DKC1* gene (3,4). Although none of the missense mutations identified so far lie directly within any of the lysine-rich regions of dyskerin, it is conceivable that they may alter the three-dimensional folding of the protein and/or lead to a disruption of interacting binding sites, both of which could influence the intracellular localization. As the mutations are scattered throughout the gene, the likelihood that these all alter the intracellular localization of dyskerin is small. Nonetheless, we considered that because there are indications that dyskerin is a multifunctional protein, there may be multiple mechanisms that play a role in the aetiology of the disease, one of which could be a mislocalization of the protein or the presence of insufficient amounts of protein at the right location. For these reasons, the missense mutations L37del, T66A, A353V and G402E in exons 3, 4, 11 and 12, respectively (3,4), were introduced into the full-length construct by *in vitro* mutagenesis and investigated in parallel with the other constructs; however, no differences were observed in the intracellular localization when compared with the wild-type construct. Together, these results indicate that mislocalization of dyskerin is unlikely to be involved in the pathogenesis of DKC.

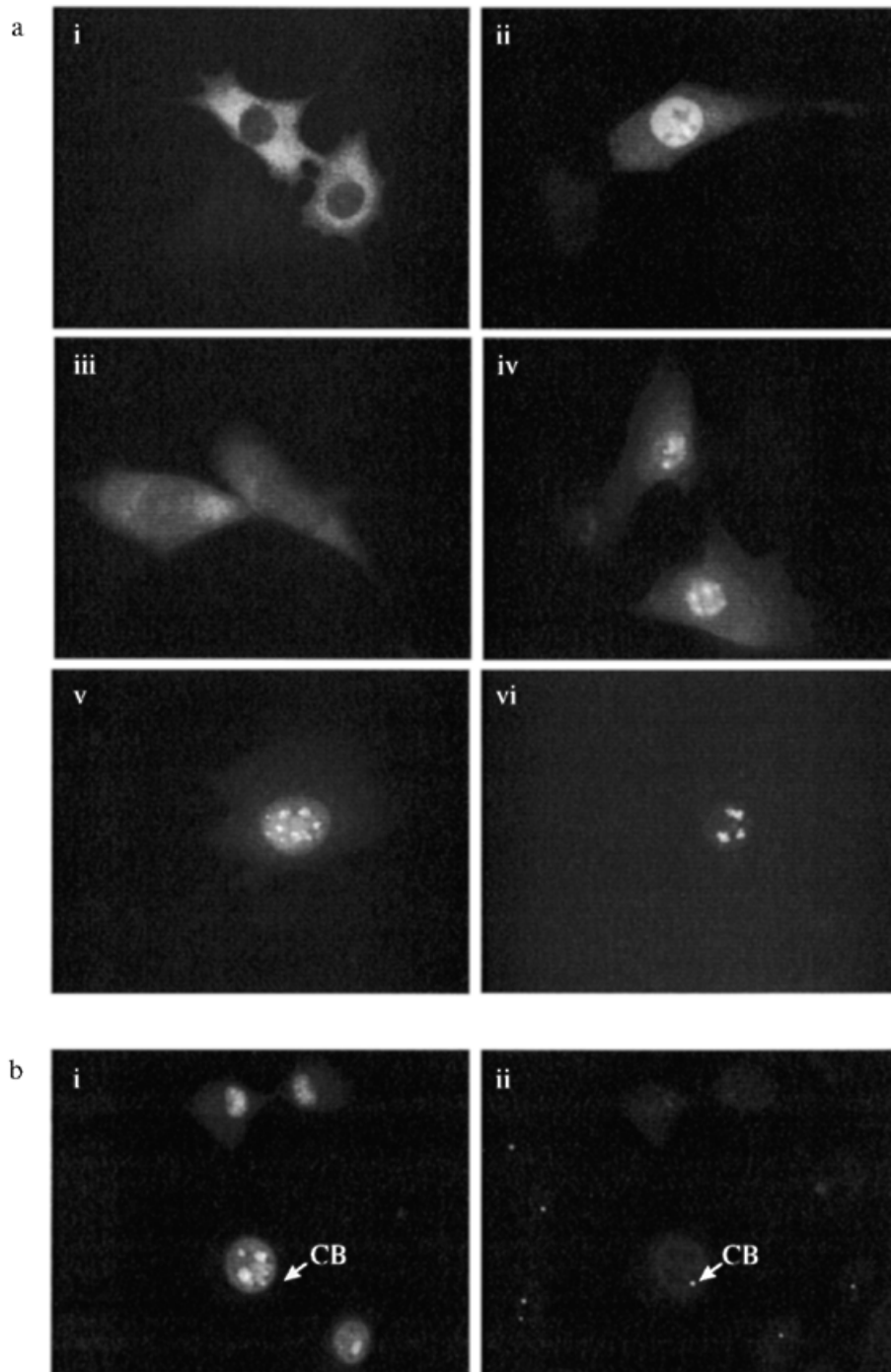
### The N- and C-terminal lysine-rich sequences co-operatively affect the nuclear and nucleolar localization of dyskerin

There are examples of proteins which require the synergistic effects of two or more bipartite NLS sequence motifs for their nuclear targeting, accumulation and retention (22). Because removal of the most C-terminal lysine-rich cluster did not visibly alter the intracellular localization of dyskerin, we decided to investigate to what extent the other putative NLS sequences are of functional significance. To this end, a number of additional deletion constructs and mutant constructs containing amino acid substitutions in the critical motifs were generated (Fig. 1).

In contrast to the full-length protein which was not detectable in the cytoplasm at any of the time points examined, deletion construct dys467 $\Delta$ 514 was expressed exclusively in the cytoplasm 2 h after microinjection (Figs 1 and 3a, i). After 4 h, dys467 $\Delta$ 514 was localizing to the cytoplasm and the nucleoplasm (Fig. 3a, ii). Unlike the full-length construct, an accumulation in the nucleoli was visible after only 6 and 8 h (Fig. 3a, iii). Similarly to the full-length construct, a translocation to the nucleoli had occurred in most cells after 24 h (Fig. 3a, vi), although cytoplasmic and nucleoplasmic fluorescence was still visible in some cells (Fig. 3a, iv and v). This indicated that the C-terminal lysine-rich clusters influence the rate of transport into the nucleus. Staining with the p80coilin antibody showed that dys467 $\Delta$ 514 co-localizes with the coiled bodies in cells where nucleolar accumulation had occurred (Fig. 3b).

Deletion construct dys446 $\Delta$ 514 localized only to the cytoplasm after 2 h (Figs 1 and 4, i). At 4, 6, 8 and 24 h after microinjection, proportionately large amounts of protein remained in the cytoplasm. In addition, an accumulation was occurring in the nucleoli after 4 h, although a nucleoplasmic localization was hardly visible at any of the investigated time points (Fig. 4, ii-iv). To determine whether this was the synergistic effect of the entire C-terminal domain or merely the effect of the sequence KRKR, these four amino acids were replaced by QAEA (dys446-QAEA-449; Fig. 1). Localization analysis of construct dys446-QAEA-449 rendered results identical to those with construct dys446 $\Delta$ 514, indicating that the sequence KRKR is primarily responsible for the nuclear import of dyskerin, but does not alone determine nucleolar localization. The seeming absence of fusion proteins dys446 $\Delta$ 514 and dys446-QAEA-449 from the nucleoplasm is unclear. Our interpretation is that the KRKR sequence contributes to a nucleoplasmic retention. Since a nucleolar localization was maintained, this suggested that there are other sequences which contribute to the nucleolar translocation of dyskerin.

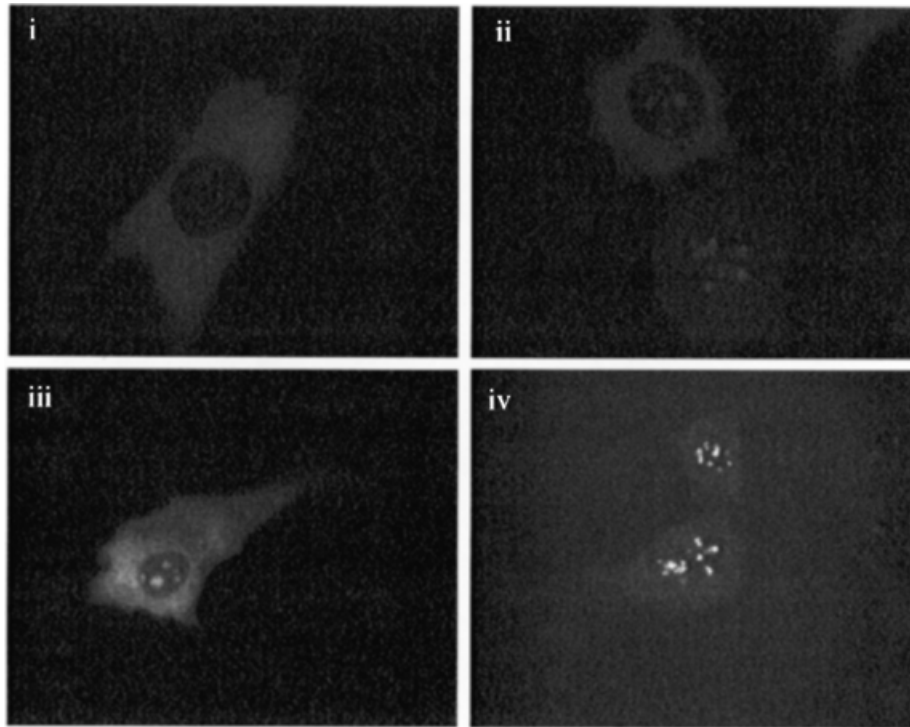
The N-terminal end of dyskerin contains the sequence KKHKKKKERKS (residues 11-21) (Fig. 1). Since complete truncation of the C-terminal lysine-rich motifs (dys-446 $\Delta$ 514) and mutation of the C-terminal KRKR sequence (dys446-QAEA-449) did not abrogate a nucleolar localization, the possibility arose that the N-terminal KKHKKKKERKS residues could be responsible for this. To delineate this sequence functionally, the deletion construct dys1 $\Delta$ 21, and the constructs dys11-NE-12, dys14-QEE-16 and dys17-QGAE-20, were investigated (Fig. 1). Compared with full-length dyskerin, neither the truncated protein nor the mutagenized proteins exhibited an altered localization pattern (Fig. 1). Together, these results suggest that although the N-terminal sequence KKHKKKKERKS influences the nucleolar localization,



**Figure 3.** (a) Representative cells showing the localization of dys467 $\Delta$ 514 tagged to EGFP at the N- or C-terminus. (i) Cytoplasmic localization observed in all cells 2 h after microinjection; (ii) cytoplasmic and nucleoplasmic localization observed in all cells 4 h after microinjection; (iii) cytoplasmic and nucleoplasmic localization with nucleolar accumulation initiating at 6 and 8 h after microinjection; (iv) cytoplasmic, nucleoplasmic and nucleolar localization observed in a proportion of cells 24 h after microinjection; (v) nucleoplasmic and nucleolar localization observed in a proportion of cells 24 h after microinjection; and (vi) sole nucleolar localization observed in a proportion of cells 24 h after microinjection. (b) (i) Nucleoplasmic and nucleolar localization of dys467 $\Delta$ 514, 24 h after microinjection showing co-localization to a coiled body; (ii) indirect immunofluorescence of the same cells stained with the p80coilin antibody verifying the position of the coiled body. CB, coiled body.

there are additional nucleolar localization signals (NuLSs). Deletion of both the N- and C-termini of dyskerin led to a complete exclusion of the fusion protein from the nucleus at all of the investigated time points (dys1 $\Delta$ 21 + 446 $\Delta$ 514; Fig. 1). Thus,

the sequences affecting the nuclear and nucleolar localization of dyskerin are contained within the N- and C-terminal lysine-rich clusters. We have shown that dyskerin harbours more than one NLS and that the NuLSs are contained within the NLSs.



**Figure 4.** Representative cells showing the localization of the constructs dys446 $\Delta$ 514 and dys446-QAEA-449. (i) Cytoplasmic localization in all cells 2 h after microinjection; (ii and iii) cytoplasmic and nucleolar localization observed in all cells 4, 6 and 8 h, and in a proportion of cells 24 h, after microinjection; and (iv) faint cytoplasmic and mainly nucleolar localization observed in a few cells 24 h after microinjection.

## DISCUSSION

Mutations in the highly conserved protein dyskerin are responsible for causing the bone marrow failure syndrome, DKC (3). Functional studies on the yeast, rat and *Drosophila* orthologues allow the prediction that dyskerin is a core component of the box H + ACA snoRNPs and that it functions in the pseudo-uridylation and cleavage of precursor rRNA (12,14). This is consistent with the present findings that dyskerin-EGFP localizes predominantly to the nucleoli when expressed in mammalian cell lines. A localization to the coiled bodies was also observed and is in agreement with the localization of rat NAP57 to both the nucleoli and the coiled bodies (8). By following a time course of expression, it became apparent that newly synthesized dyskerin initially localizes to the nucleoplasm, followed by a sequential translocation to the nucleoli and the coiled bodies. A co-localization to the coiled bodies occurred only in those cells where dyskerin had already accumulated in the nucleoli. Nucleolar transport kinetics have not been demonstrated for NAP57, although a similar temporal lag in the targeting to the nucleoli and coiled bodies was observed for the interacting protein of NAP57, Nopp140 (19). This led to the conclusion that the path of Nopp140 to the coiled bodies leads through the nucleoli (19), and provides a feasible explanation for our observations. There is evidence that Nopp140 shuttles on curvilinear tracks between the cytoplasm and the nucleus (17), as well as between the nucleoplasm, the nucleolus and the coiled bodies, and that NAP57 is escorted along the same pathways (19). As direct evidence is still lacking, it will be important to

determine whether dyskerin is dynamically associated with these structures.

The primary sequence of dyskerin harbours lysine-rich NLSs at the N- and C-termini. The C-terminal motif KRKR(X)<sub>17</sub>KKEKKKSKKDKKAK(X)<sub>17</sub>KKKKKKKKAKEVELVSE in particular contains clusters of overlapping NLSs which most closely resemble the C-terminal motifs described for the Treacher Collins syndrome protein treacle (29). Treacle-GFP fusions lacking all or part of the C-terminal NLSs, respectively, fail to localize or localize inefficiently to the nucleoli (30,31). In our study, deletion of the most C-terminal lysine-rich cluster which mimics the patient deletion (dys493 $\Delta$ 514; Fig. 1) did not alter discernibly the localization of dyskerin to the nucleolus and the coiled bodies. Considering that the endogenous expression of dyskerin is expected to differ quantitatively and qualitatively from the exogenous expression of the EGFP fusions, this result should be considered with caution. In the patient, the open reading frame is predicted to continue into the last intron, which results in a replacement of the lysine-rich cluster (DSDTTKKKKKKKKAKEVELVSE) by 12 new amino acids (VCGNTFRFLGLA) (5) which were not included in construct dys493 $\Delta$ 514. Since the seven extra amino acids (GPGSIAT) introduced by the multiple cloning site in C-terminally tagged dyskerin did not affect localization, and no differences in localization were observed between the N- and C-terminally tagged constructs, this indicates that the signals responsible for targeting dyskerin to the nucleus override the effect of the foreign amino acids. In this context, it is interesting that although dyskerin is highly

conserved between the species, the C-terminal motifs are least well conserved (3). Further, a polymorphism leading to an insertion of an additional lysine in the stretch of eight lysines is tolerated (4). We conclude that the NLSs exert a dominant effect over the foreign amino acids in targeting dyskerin to the nucleus. It was not surprising, therefore, that none of the missense mutations mimicking patient mutations showed an altered localization pattern. Although our present results basically rule out that DKC is caused by a mislocalization of dyskerin, it would still be interesting to examine the expression and localization of endogenous dyskerin with anti-dyskerin antibodies in DKC patient cell lines. For this and other reasons, we are in the process of characterizing monoclonal antibodies. There are examples of diseases, such as spinal muscular atrophy (32,33) and acute promyelocytic leukaemia (34,35), where mutations do not cause the proteins to mislocalize, but rather cause a disruption in the organization of the nucleus and the subnuclear bodies. This should also be investigated in DKC patient cell lines.

The transport of proteins into the nucleus is a complex process requiring the co-operative effects of several factors and signal sequences which regulate import, retention and export. An extended analysis of the C-terminal NLS motifs in dyskerin indicated that full nuclear accumulation is accomplished by the additive effects of the N- and C-terminal NLS sequence clusters. Whereas deletion of the most C-terminal lysine-rich cluster (dys493 $\Delta$ 514) had no visible effect on the intracellular localization of dyskerin, deletion of both lysine-rich clusters (dys467 $\Delta$ 514) unambiguously delayed nuclear accumulation. This indicates that the most C-terminal lysine-rich clusters affect the rate of nuclear transport. A role for phosphorylation in increasing the rate of nuclear uptake was documented originally for the SV40 large T antigen (36–38), and was shown to regulate nuclear uptake of the transcription factor Pho4 in yeast both positively and negatively (39). The rate of nuclear localization of human parathyroid hormone-related protein is controlled in a phosphorylation- and cell cycle-dependent manner (40), and this has been proposed as a common mechanism for regulating the interaction of nuclear proteins with nuclear import and export factors (41). Since the C-terminal region of dyskerin contains a number of putative phosphorylation sites (3), it is possible that the domains deleted in dys467 $\Delta$ 514 are not themselves nuclear targeting signals, but that they co-ordinate the rate of nuclear translocation by providing phosphorylation sites. Similarly, the delayed nuclear uptake of deletion constructs of the Fanconi anaemia group A protein, FANCA, was attributed to the loss of phosphorylation sites and was proposed to cause the mixed pattern of cytoplasmic and/or nuclear localization (42). Since in dyskerin the putative phosphorylation sites are not confined to the C-terminus, this could also be the reason why a mixed localization pattern was observed for full-length dyskerin and could explain why nucleolar accumulation occurred more rapidly in some cells than in others. Conceivably, the extent to which dyskerin is phosphorylated depends on the stage of the cell cycle. It will be interesting to study the kinetics of nuclear transport of dyskerin by using synchronized cells.

The cytoplasmic localization pattern of dys446 $\Delta$ 514 and dys446-QAEA-449 indicated that the C-terminal KRKR sequence is primarily responsible for the nuclear uptake of dyskerin. Simultaneously it seems to regulate nucleoplasmic

retention, while only secondarily influencing a translocation into the nucleoli. There appears to be an NuLS in the N-terminal sequence KKHKKKKKERK. Since a nucleolar localization still occurred when the N-terminal sequence KKHKKKKKERK was deleted or modified (dys11-DE-12, dys14-EE-16 and dys17-QGAE-20) and deletion of both the N- and C-terminal motifs completely abolished the nucleoplasmic and nucleolar accumulation of dyskerin (dys1 $\Delta$ 21 + 446 $\Delta$ 514), this shows that the intranuclear transport of dyskerin is a complicated process which is affected additively by sequences in the N- and C-termini. This agrees with findings that although a large number of nuclear and nucleolar proteins in numerous species have been characterized, no universal consensus for NLSs exists, and even less so for NuLSs. The nucleolar localization of nucleolin, for example, requires RNA-binding domains and a C-terminal stretch of amino acids rich in glycine and arginine (24). Contrary to this, a C-terminal acidic domain and a DNA-binding region are necessary for the nucleolar localization of the transcription factor UBF (43). In the case of treacle, the C-terminal sequence KRKKDKEKKEKKKKAKK contributes to the nucleolar localization (30,31). The results for dyskerin confirm and extend the notion that NuLSs are unique to specific proteins and in some cases represent extended NLSs.

At present, it is not known how the various mutations affect the functions of dyskerin to cause DKC. Even though the *DKC1* transcript is expressed ubiquitously (3) and dyskerin localizes to the nucleolus which is a ubiquitous organelle, the DKC phenotype shows specificity in that it is mainly the rapidly dividing cells which are clinically affected in DKC patients. It will be a difficult task to unravel the mechanisms that lead to the disease, especially since there is such a range of mutations scattered across the gene. With the exception of the exon 15 deletion, a putative splice site mutation in intron 2 and a single amino acid deletion in exon 3, DKC is caused by missense mutations in exons 1, 3, 4, 10, 11 and 12 of the *DKC1* gene (3–5). The mutation hot spots in exons 3, 4 and 11, however, imply that they contain important functional residues, although they do not lie within the TruB and NLS domains of the primary sequence. No whole gene deletions, frameshift or stop codon mutations have been identified so far, suggesting that these would be lethal. This is compatible with findings that total loss-of-function mutations are non-viable both in yeast and in *Drosophila* (9,11,12) and underlines the essential function of dyskerin. The mutations detected in DKC patients can therefore be expected to alter one of the many functions of dyskerin more subtly.

Truncation of the C-terminal KKE/KKD repeat domain in yeast Cbf5p leads to a delay at the G<sub>2</sub>/M phase of the cell cycle (9). Cell lines from DKC patient skin fibroblasts and from peripheral blood lymphocytes exhibit increased doubling times and fewer cell divisions. In many cases, this is accompanied by an increased chromosomal instability (1,44). The rapid senescence of DKC cells *in vitro* is reminiscent of a DNA damage-induced cell cycle arrest. Although this may only be related indirectly to a dysfunction in programmed cell death, the apoptotic cells observed in the ovaries and degenerating embryos of *Drosophila* mutants expressing reduced levels of Nop60B adds credibility to such speculations (12). Further, there could be a connection between the untimely cell death of DKC cell lines and findings that human telomerase shows some homology to the box H + ACA snoRNAs (45) with which dyskerin is thought to interact.

In view of the links that have been made between the nucleolus, cell cycle regulation, ageing and disease (46,47), it is intriguing that DKC patients exhibit some features of premature ageing (48). To address these functional issues more directly, we are in the process of generating a murine model for the disease.

## MATERIALS AND METHODS

### Generation of dyskerin-EGFP fusion constructs

*DKC1* cDNA fragments were PCR amplified with forward and reverse primers containing *Bgl*III and *Xma*I restriction sites, respectively. To reduce the error rate, 15 cycles and an Amplitaq:Pfu enzyme mix (8:1) were used. The following primer combinations were used for amplification of inserts for cloning into the pEGFP-C1 vector (Clontech, Heidelberg, Germany):

construct dys1-514:  
X101cBF3 (GGAAGATCTGCGGATGCGGAAGTA)  
and X101cXR1 (TCCCCCGGGCTACTCAGAAACC);  
construct dys493Δ514:  
X101cBF3 (GGAAGATCTGCGGATGCGGAAGTA)  
and X101cXR4 (TCCCCCGGGCTACCCATCTCCAGG);  
construct dys467Δ514:  
X101cBF3 (GGAAGATCTGCGGATGCGGAAGTA)  
and X101cXR5 (TCCCCCGGGCTAGATCAACTGAGG);  
construct dys1Δ21 + 446Δ514:  
X101cBF7 (GGAAGATCTTTGCCAGAAGAAGAT)  
and X101cXR7 (TCCCCCGGGCTACGCAGTTTTTGC).

The following primer combinations were used for amplification of inserts for cloning into the pEGFP-N3 vector (Clontech):

construct dys1-514:  
X101nBF3 (GGAAGATCTATGGCGGATGCGGAAGTA)  
and X101nXR1 (TCCCCCGGGCCCTCAGAAACCAATTC);  
construct dys493Δ514:  
X101nBF3 (GGAAGATCTATGGCGGATGCGGAAGTA)  
and X101nXR4 (TCCCCCGGGCCCCATCTCCAGG);  
construct dys467Δ514:  
X101nBF3 (GGAAGATCTATGGCGGATGCGGAAGTA)  
and X101nXR5 (TCCCCCGGGCCGATCAACTGAGG);  
construct dys1Δ21 + 446Δ514:  
X101nBF7 (GGAAGATCTATGTTGCCAGAAGAAGAT)  
and X101nXR7 (TCCCCCGGGCCCCGCAGTTTTTGC).

Constructs dys446Δ514 and dys1Δ21 were generated by digesting clone dys1-514 with *Bgl*III, *Pst*I and *Xcm*I, and religating the respective fragments into the EGFP-N3 vector.

Amino acid substitutions were introduced into N-terminally tagged clone dys1-514 by using the QuikChange site-directed mutagenesis kit (Stratagene, Amsterdam, The Netherlands). The reactions were carried out according to the manufacturer's instructions. The following primers were used for *in vitro* mutagenesis:

construct dys446-QAEA-449:  
dys446-QAEA-449f (GCAGCAAAACTGCGCAAGCGG-  
AAGCGGAGAGTGAGAGTG)  
and dys446-QAEA-449r (CACTCTCACTCTCCGCTTCCGCT-  
TGCGCAGTTTTGCTGC);  
construct dys11-NE-12:  
dys11-NE-12f (GTAATTATTTTGCCAAACGAACATAA-  
GAAGAAAAG)

and dys11-NE-12r (CTTTTTCTTCTTATGTTCGTTTGGC-  
AAAATAATTAC);  
construct dys14-QEE-16:  
dys14-QEE-16f (GCCAAAGAAACATCAAGAAGAAAAG-  
GAGCGGAAG)  
and dys14-QEE-16r (CTTCCGCTCCTTTTCTTCTTGATG-  
TTTCTTTGGC);  
construct dys17-QGAE-20:  
dys17-QGAE-20f (GAAACATAAGAAGAAACAAGGAGC-  
GGAGTCATTGCCAGAAGAAG)  
and dys17-QGAE-20r (CTTCTTCTGGCAATGACTCCGCT-  
CCTTGTTTTCTTC-TTATGTTTC).

The following primers were used for introducing patient missense mutations:

L37del-f (CAACACGCTGAAGAATTTATCAAACCTGA-  
ATCCAAAG)  
and L37del-r (CTTTGGATTTCAGGTTTGATAAATTCCTTC-  
AGCGTGTTG);  
T66A-f (GCTGAATGTAAGGGCAACACACTATAACAC)  
and T66A-r (GTGTATAGTGTGTTGCCCTTACATTCAGC);  
A353V-f (GCATTAATGACCACAGTGGTCATCTCTACC-  
TGCG)  
and A353V-r (CGCAGGTAGAGATGACCACTGTGGTCA-  
TTAATGC);  
G402R-f (CCTTCTGGACAAGCATAGGAAGCCCACAG-  
ACAG)  
and G402R-r (CTGTCTGTGGGCTTCCATGCTTGTCCA-  
GAAGG).

All constructs were sequenced to confirm that the inserts were in-frame, to exclude unwanted mutations and to check for introduced mutations.

### Cell culture and transfections

HeLa and NIH-3T3 cells were cultured in Dulbecco's modified Eagle's medium and Vero cells in minimal essential medium (MEM) supplemented with 10% fetal calf serum, 10 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin. Cells were grown in a humidified incubator with 5% CO<sub>2</sub> at 37°C. For the transfections, the cells were seeded out onto glass coverslips (1.5 cm diameter) at a density of 4 × 10<sup>5</sup> per well (3.5 cm diameter) and allowed to settle overnight. NIH-3T3 and Vero cells were transfected with 15 µl of Lipofectamine (Gibco BRL, Karlsruhe, Germany) and 2.5 µg of DNA. On average, the transfection efficiencies were >25% for NIH-3T3 cells and <25% for the Vero cells. HeLa cells were transfected with 5 µl of Effectene transfection reagent (Qiagen, Hilden, Germany), 1.6 µl of enhancer and 300 ng of DNA. Transfection efficiencies of up to 80% were achieved for HeLa cells.

### Microinjections, fluorescence microscopy and imaging

Cells were grown overnight in 3.5 cm diameter glass bottom dishes (MatTek, Ashland, MA). Immediately prior to the microinjections, the culture medium was replaced with medium containing bicarbonate (0.85 g/l) and 30 mM HEPES to maintain the pH. The microinjection needles were pulled from borosilicate glass capillaries GC 100F-10 (Clark Electromedical Instruments, Pangbourne, UK) using a Flaming/Brown Micropipette Puller (Model P-97; Sutter Instruments, Novato, CA). The cells were injected with 10–20 ng of DNA on a stage heated to 37°C using the Micromanipulator 5171 and Transjector 5246 (Eppendorf, Hamburg, Germany) (49).



After incubation in MEM at 37°C, the cells were imaged at the indicated time points. Microscopy and imaging was performed on a Leica DM IRBE with a 63× NA1.4 PL Apo objective (Leica, Bensheim, Germany), a GFP filter set and a Hamamatsu CCD camera C4742-95. The images were processed on Macintosh computers using Openlab 2.0 and Adobe Photoshop 5.0.

### Immunofluorescence

Transfected and microinjected cells were washed twice with phosphate-buffered saline (PBS), fixed for 20 min with 3.7% formaldehyde and permeabilized for 15 min with 0.25% Triton-X. Cells were blocked in 3% bovine serum albumin (BSA) (in PBS) for 1 h at room temperature. Primary antibodies (diluted in 3% BSA) were incubated for 2 h at 37°C in a humid chamber. The cells were then washed three times for 5 min in PBS, 0.01% Tween-20. Secondary antibodies (diluted in 3% BSA, 0.01% 4',6-diamidino-2-phenylindole) were incubated for 1 h at room temperature. The cells were again washed three times for 5 min in PBS, 0.01% Tween-20, air dried, sealed and viewed with a Zeiss microscope. Primary antibodies used were B23 (diluted 1:50; Santa Cruz Biotechnologies, Santa Cruz, CA) and p80coilin (diluted 1:350; gift from Angus Lamond, University of Dundee, UK). Secondary antibodies were the anti-goat and anti-rabbit IgG Cy3 conjugates (diluted 1:300; Sigma, Deisenhofen, Germany).

### ACKNOWLEDGEMENTS

We thank Angus Lamond (Dundee, UK) for generously providing the p80coilin antibody, Z. Sedlacek and I. Krebs for critical reading of the manuscript, and Eppendorf-Netheler-Hinz GmbH (Hamburg, Germany), Improvisation (Coventry, UK), Leica Microsystems (Germany), Hamamatsu Photonics (Hamamatsu City, Japan), Olympus Optical Co. Europa GmbH (Hamburg, Germany) and T.I.L.L. Photonics GmbH (Martinsried, Germany) for providing the EMBL with the Advanced Light Microscopy Facility. This work was supported by grants from the Deutsche Forschungsgemeinschaft (DFG) and from the European Community (EU) Genome Analysis Program.

### REFERENCES

- Dokal, I. and Luzzatto, L. (1994) Dyskeratosis congenita is a chromosomal instability disorder. *Leukem. Lymphoma*, **15**, 1–7.
- Knight, S., Vulliamy, T., Copplestone, A., Gluckman, E., Mason, P. and Dokal, I. (1998) Dyskeratosis Congenita (DC) Registry: identification of new features of DC. *Br. J. Haematol.*, **103**, 990–996.
- Heiss, N.S., Knight, S.W., Vulliamy, T.J., Klauck S.M., Wiemann, S., Mason, P.J., Poustka, A. and Dokal, I. (1998) X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nature Genet.*, **19**, 32–38.
- Knight, S.W., Heiss, N.S., Vulliamy, T.J., Greschner, S., Stavrides G., Pai, G.S., Lestringant, G., Varma, N., Mason, P.J., Dokal, I. and Poustka, A. (1999) X-linked dyskeratosis congenita is predominantly caused by missense mutations in the *DKC1* gene. *Am. J. Hum. Genet.*, **65**, 50–58.
- Vulliamy, T.J., Knight, S.W., Heiss, N.S., Smith, O.P., Poustka, A., Dokal, I. and Mason, P.J. (1999) Dyskeratosis congenita caused by a 3' deletion: germ line and somatic mosaicism in a female carrier. *Blood*, **94**, 1254–1260.
- Devriendt, K., Matthijs, G., Legius, E., Schollen, E., Blockmans, D., van Geet, C., Degreef, H., Cassiman, J.-J. and Fryns, J.-P. (1997) Skewed X-chromosome inactivation in female carriers of dyskeratosis congenita. *Am. J. Hum. Genet.*, **60**, 581–587.
- Vulliamy, T.J., Knight, S.W., Dokal, I. and Mason, P.J. (1997) Skewed X-inactivation in carriers of X-linked dyskeratosis congenita. *Blood*, **90**, 2213–2216.
- Meier, U.T. and Blobel, G. (1994) NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria. *J. Cell Biol.*, **127**, 1505–1514.
- Jiang, W., Middleton, K., Yoon, H.-J., Fouquet, C. and Carbon, J. (1993) An essential yeast protein, CBF5p, binds *in vitro* to centromeres and microtubules. *Mol. Cell. Biol.*, **13**, 4884–4893.
- Cadwell, C., Yoon, H.-J., Zebrarjadian, Y. and Carbon, J. (1997) The yeast nucleolar protein Cbf5p is involved in rRNA biosynthesis and interacts genetically with the RNA polymerase I transcription factor RRN3. *Mol. Cell. Biol.*, **17**, 6175–6183.
- Phillips, B., Billin, A.N., Cadwell, C., Buchholz, R., Erickson, C., Merriam, J.R., Carbon, J. and Poole, S.J. (1998) The *Nop60B* gene of *Drosophila* encodes an essential nucleolar protein that functions in yeast. *Mol. Gen. Genet.*, **260**, 20–29.
- Giordano, E., Peluso, I., Senger, S. and Furia, M. (1999) *Minify*, a *Drosophila* gene required for ribosome biogenesis. *J. Cell Biol.*, **144**, 1123–1133.
- Ni, J., Tien, A.L. and Fournier, M.J. (1997) Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell*, **89**, 565–573.
- Lafontaine, D.L.J., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. and Tollervey, D. (1998) The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev.*, **12**, 527–537.
- Henras, A., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gelugne, J.P. and Caizergues-Ferrer, M. (1998) Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J.*, **17**, 7078–7090.
- Koonin, E.V. (1996) Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. *Nucleic Acids Res.*, **24**, 2411–2415.
- Meier, U.T. and Blobel, G. (1992) Nopp140 shuttles on tracks between nucleolus and cytoplasm. *Cell*, **70**, 127–138.
- Lamond, A.I. and Earnshaw, W.C. (1998) Structure and function in the nucleus. *Science*, **280**, 547–553.
- Isaac, C., Yang, Y. and Meier, U.T. (1998) Nopp140 functions as a molecular link between the nucleolus and the coiled bodies. *J. Cell Biol.*, **142**, 319–329.
- Matera, A.G. and Frey, M.R. (1998) Coiled bodies and gems: Janus or gemini? *Am. J. Hum. Genet.*, **63**, 317–321.
- Pederson, T. (1998) The plurifunctional nucleolus. *Nucleic Acids Res.*, **26**, 3871–3876.
- Dingwall, C. and Laskey, R.A. (1991) Nuclear targeting sequences—a consensus? *Trends Biochem. Sci.*, **16**, 478–481.
- Chelsky, D., Ralph, R. and Jonak, G. (1989) Sequence requirements for synthetic peptide-mediated translocation to the nucleus. *Mol. Cell. Biol.*, **9**, 2487–2492.
- Schmidt-Zachmann, M.S. and Nigg, E.A. (1993) Protein localization to the nucleolus: a search for targeting domains in nucleolin. *J. Cell Sci.*, **105**, 799–806.
- Quaye, I.K., Toku, S. and Tanaka, T. (1996) Sequence requirement for nucleolar localization of rat ribosomal protein L31. *Eur. J. Cell Biol.*, **69**, 151–155.
- Görllich, D. and Mattaj, I.W. (1996) Nucleocytoplasmic transport. *Science*, **271**, 1513–1518.
- Englmeier, L., Olivo, J.C. and Mattaj, I.W. (1999) Receptor-mediated substrate translocation through the nuclear pore complex without nucleotide triphosphate hydrolysis. *Curr. Biol.*, **9**, 30–41.
- Makkerh, J.P.S., Dingwall, C. and Laskey, R.A. (1996) Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Curr. Biol.*, **6**, 1025–1027.
- Dixon, J., Edwards, S.J., Gladwin, A.J., Dixon, J., Loftus, S.K., Bonner, C.A., Koprivnikar, K. and Wasmuth, J.J. (1996) Positional cloning of a gene involved in the pathogenesis of Treacher Collins syndrome. *Nature Genet.*, **12**, 130–136.
- Marsh, K.L., Dixon, J. and Dixon, M.J. (1998) Mutations in the Treacher Collins syndrome gene lead to mislocalization of the nucleolar protein treacle. *Hum. Mol. Genet.*, **7**, 1795–1800.
- Winokur, S.T. and Shiang, R. (1998) The Treacher Collins syndrome (*TCOF1*) gene product, treacle, is targeted to the nucleolus by signals in its C-terminus. *Hum. Mol. Genet.*, **7**, 1947–1952.
- Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Bulet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P. and Zeviani, M. (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*, **80**, 155–165.

33. Liu, Q. and Dreyfuss, G. (1996) A novel nuclear structure containing the survival of motor neurons protein. *EMBO J.*, **15**, 3555–3565.
34. de The, H., Lavau, C., Marchio, A., Chomienne, C., Degos, L. and Dejean, A. (1991) The PML–RAR alpha fusion mRNA generated by the t(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. *Cell*, **66**, 675–684.
35. Hodges, M., Tissot, C., Howe, K., Grimwade, D. and Freemont, P.S., (1998) Structure, organization, and dynamics of promyelocytic leukemia protein nuclear bodies. *Am. J. Hum. Genet.*, **63**, 297–304.
36. Jans, D.A. and Jans, P. (1994) Negative charge at the casein kinase II site flanking the nuclear localization signal of the SV40 large T-antigen is mechanistically important for enhanced nuclear import. *Oncogene*, **9**, 2961–2968.
37. Hubner, S., Xiao, C.Y. and Jans, D.A. (1997) The protein kinase CK2 site (Ser111/112) enhances recognition of the simian virus 40 large T-antigen nuclear localization sequence by importin. *J. Biol. Chem.*, **272**, 17191–17195.
38. Xiao, C.Y., Hubner, S. and Jans, D.A. (1997) SV40 large tumor antigen nuclear import is regulated by the double-stranded DNA-dependent protein kinase site (serine 120) flanking the nuclear localization sequence. *Biol. Chem.*, **272**, 22191–22198.
39. Komeili, A. and O’Shea, E.K. (1999) Roles of phosphorylation sites in regulating activity of the transcription factor Pho4. *Science*, **284**, 977–980.
40. Lam, M.H., House, C.M., Tiganis, T., Mitchelhill, K.I., Sarcevic, B., Cures, A., Ramsay, R., Kemp, B.E., Martin, T.J. and Gillespie M.T. (1999) Phosphorylation at the cyclin-dependent kinases site (Thr85) of parathyroid hormone-related protein negatively regulates its nuclear localization. *J. Biol. Chem.*, **274**, 18559–18566.
41. Hood, J.K. and Silver, P.A. (1999) In or out? Regulating nuclear transport. *Curr. Opin. Cell. Biol.*, **11**, 241–247.
42. Lightfoot, J., Alon, N., Bosnoyan-Collins, L. and Buchwald, M. (1999) Characterization of regions functional in the nuclear localization of the Fanconi anemia group A protein. *Hum. Mol. Genet.*, **8**, 1007–1015.
43. Maeda, Y., Hisatake, K., Kondo, T., Hanada, K., Song, C.Z., Nishimura, T. and Muramatsu, M. (1992) Mouse rRNA gene transcription factor mUBF requires both HMG-box1 and an acidic tail for nucleolar accumulation: molecular analysis of the nucleolar targeting mechanism. *EMBO J.*, **11**, 3695–3704.
44. Coulthard, S., Chase, A., Pickard, J., Goldman, J. and Dokal, I. (1998) Chromosomal breakage analysis in dyskeratosis congenita peripheral blood lymphocytes. *Br. J. Haematol.*, **102**, 1162–1164.
45. Mitchell, J.R., Cheng, J. and Collins, K. (1999) A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3’ end. *Mol. Cell. Biol.*, **19**, 567–576.
46. Guarente, L. (1997) Link between ageing and the nucleolus. *Genes Dev.*, **11**, 2449–2455.
47. Garcia, S.N. and Pillus, L. (1999) Net results of nucleolar dynamics. *Cell*, **97**, 825–828.
48. Luzzatto, L. and Karadimitris, A. (1998) Dyskeratosis and ribosomal rebellion. *Nature Genet.*, **19**, 6–7.
49. Pepperkok, R., Scheel, J., Horstmann, H., Hauri, H.P., Griffiths, G. and Kreis, T.E. (1993) Beta-COP is essential for biosynthetic membrane transport from the endoplasmic reticulum to the Golgi complex *in vivo*. *Cell*, **74**, 71–82.