Pathogenic NAP57 mutations decrease ribonucleoprotein assembly in dyskeratosis congenita

Petar N. Grozdanov1,†, Narcis Fernandez-Fuentes2,3,‡, Andras Fiser2,3 and U. Thomas Meier1,*

1Department of Anatomy and Structural Biology, 2Department of Systems and Computational Biology and 3Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA

Received July 20, 2009; Revised August 13, 2009; Accepted August 26, 2009

X-linked dyskeratosis congenita (DC) is a rare bone marrow failure syndrome caused by mostly missense mutations in the pseudouridine synthase NAP57 (dyskerin/Cbf5). As part of H/ACA ribonucleoproteins (RNPs), NAP57 is important for the biogenesis of ribosomes, spliceosomal small nuclear RNPs, microRNAs and the telomerase RNP. DC mutations concentrate in the N- and C-termini of NAP57 but not in its central catalytic domain raising questions as to their impact. We demonstrate that the N- and C-termini together form the binding surface for the H/ACA RNP assembly factor SHQ1 and that DC mutations modulate the interaction between the two proteins. Pinpointing impaired interaction between NAP57 and SHQ1 as a potential molecular basis for X-linked DC has implications for therapeutic approaches, e.g. by targeting the NAP57–SHQ1 interface with small molecules.

INTRODUCTION

Dyskeratosis congenita (DC) is a rare inherited bone marrow failure syndrome characterized by the cutaneous triad of nail dystrophy, abnormal skin pigmentation and mucosal leukopla-kia, causing premature death due to aplastic anemia (1). Although in many cases the genes affected remain to be identified, the predominant X-linked form is caused by mutations in the pseudouridine synthase, NAP57 (aka dyskerin or Cbf5 in yeast and Cbf5 in archaea), of H/ACA ribonucleoproteins (RNPs) (2,3). The autosomal dominant and recessive forms trace to mutations in two additional H/ACA RNP core proteins, NOP10 and NHP2, in telomerase reverse transcriptase and RNA, and in the telomeric shelterin component TIN2. Given that telomerase belongs to the family of H/ACA RNPs and all DC patients exhibit shortened telomeres in their peripheral blood cells, DC is considered a telomere deficiency (1). Nevertheless, it is contentious to what degree impairment of the functions of the other 150 or so H/ACA RNPs, e.g. ribosomal RNA and spliceosomal small nuclear RNA modification, contribute to the pathological features of DC (4,5). Importantly, the molecular impact of none of the identified mutations in any of the affected genes has been elucidated. Here we demonstrate that the mutations in NAP57 modulate its interaction with SHQ1, one of the assembly factors of H/ACA RNPs.

H/ACA RNPs consist of two sets of four core proteins, each set assembling around one of two hairpins of a function determining H/ACA RNA (6). In most cases, H/ACA RNAs guide site-directed pseudouridylation of target RNAs by base pairing (7). Other H/ACA RNAs, such as telomerase RNA, which ends in an H/ACA domain, lack complementarity to any stable RNAs (8). NAP57 together with NOP10 and NHP2 forms a core trimer that is essential for the stability of all H/ACA RNPs. A fourth H/ACA core protein, GAR1, interacts with NAP57 independently (9). A hairpin of an H/ACA RNA stretches across the relatively planar surface of the core trimer anchored with its 3'-end in the PUA domain of NAP57 and with its terminal loop in NHP2 (10). The mostly missense mutations in NAP57 identified in X-linked DC patients concentrate in its N-terminal (NT) and C-terminal (CT) domains mostly avoiding the central catalytic domain (Fig. 2A) (11). In models of the three-dimensional structure of human NAP57, the mutated domains wrap around each

†Present address: Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, 3601 4th Street, Lubbock, TX 79430, USA.
‡Present address: Leeds Institute of Molecular Medicine, Section of Experimental Therapeutics, University of Leeds, St James’s University Hospital, Wellcome Trust Brenner Building, Leeds LS9 7TF, UK.

© The Author 2009. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oxfordjournals.org
RESULTS

We previously demonstrated that SHQ1 binds NAP57 making it a candidate for a factor interacting with the DC mutation cluster of NAP57 (13). However, testing this hypothesis required the fusion of the NT and CT domains of NAP57 forming the mutation cluster while excising its catalytic domain. To achieve this goal without disrupting the tertiary structure of the mutation cluster, we modeled the structure of human NAP57 based on those of the crystal structures of archaeal H/ACA RNPs. Archaeal H/ACA RNPs differ from mammalian ones in that their NHP2 ortholog, L7Ae, independently recognizes archaeal H/ACA RNAs, whereas NHP2 together with the other H/ACA core proteins can form a protein only complex (9,10). Therefore, we modeled this human complex by docking NHP2, modeled according to the structure of the related 15.5 K protein (14), onto the NOP10–NAP57–GAR1 complex, as described in Materials and Methods (Fig. 1A). In the model of human NAP57 derived from this complex, we identified two helices at the NT and CT of the catalytic domain that were in close proximity to each other. Specifically, fusion of the residues G90 and P288, which are only 5.05 Å apart, would generate a new protein, NAP57\textsubscript{cat}, with minimal impact on the tertiary structure of its CT PUA domain benefitting from the fact that Gly-Pro motifs are commonly found in naturally occurring kinked helices (Fig. 1B) (15). Thus, we designed and computationally assessed the feasibility (see Materials and Methods) of a NAP57 construct, NAP57\textsubscript{cat}, that lacked its catalytic domain but retained most DC mutation sites (Figs 1B and 2A).

To experimentally address if the NAP57\textsubscript{cat} construct folded properly, it was expressed in bacteria as a fusion with the maltose binding protein (MBP, Fig. 2A) and tested in amylose resin pull-down assays (13). Specifically, we investigated if MBP-NAP57\textsubscript{cat} (and additional constructs, Fig. 2A) retained the ability of full-length NAP57 to interact with recombinant SHQ1 (Fig. 2B). All constructs containing the CT domain of NAP57 bound SHQ1 (Fig. 2B, lanes 8–10), whereas the catalytic domain (cat, lane 7) or the NT alone did not (Fig. 2C, lane 4). To further explore the contribution of the Δcat NT to the interaction with SHQ1, binding of SHQ1 to Δcat and to the CT alone was compared at increasing salt concentrations (Fig. 2D). Binding was clearly enhanced by the presence of the NT, which was particularly evident at 210 mM salt (Fig. 2D, compare lanes 5 and 6, inset). This result was corroborated when investigated in the context of the cell nucleus using an in vivo nuclear tethering assay (16). Specifically, when tethered (by a Lac repressor fusion domain, Laci) to Lac operator arrays integrated in the genome, Δcat recruited endogenous SHQ1 efficiently (Fig. 2E, 84% of cells, $n = 50$), whereas the CT alone failed to do so (Fig. 2F, 79% of cells, $n = 68$). These data indicated that the NT folded properly around the CT in our NAP57\textsubscript{cat} construct and that both domains concentrating most DC mutation sites were required for full interaction with SHQ1. In conclusion, SHQ1 bound to the DC mutation cluster.

SHQ1 consists of an HSP20-like NT CS domain and a unique CT (Fig. 3A) (13,17,18). We tested if SHQ1 lacking its CS domain, SHQ1\textsubscript{ΔCS}, bound to NAP57\textsubscript{Δcat}, because it was this domain that interacted with full-length NAP57 (13). Indeed, SHQ1\textsubscript{ΔCS} interacted with NAP57\textsubscript{Δcat} (Fig. 3B), whereas the CS domain alone failed to do so (Fig. 3C) further supporting proper folding of the NAP57\textsubscript{Δcat} construct. Moreover, recruitment of endogenous SHQ1 to sites of NAP57\textsubscript{Δcat} concentration in the nucleus (Fig. 2E) was specific, because Δcat failed to recruit another H/ACA RNP assembly factor, NAF1 (Fig. 3D), and because the catalytic domain alone did not recruit SHQ1 (Fig. 3E). Finally, when cotransfected with NAP57\textsubscript{Δcat}, the normally cytoplasmic SHQ1-Laci (13) was ferried quantitatively into the nucleus [presumably by the abundance of nuclear localization signals in the NAP57\textsubscript{Δcat} domains (19)] where it concentrated together with Δcat in nucleoli and at Lac operator sites (Fig. 3F). Together, these data provide further evidence for a specific interaction between SHQ1\textsubscript{ΔCS} and NAP57\textsubscript{Δcat}.

Our data showed that SHQ1 bound to the NAP57 domains harboring most of the mutation sites identified in X-linked DC patients (Figs 1–3). Therefore, we tested whether DC mutations interfered with SHQ1 binding. The predominant mutation in X-linked DC is a mutation of alanine 353 to valine (A353V), accounting for ~40% of all cases (11). Owing to the close relatedness of alanine and valine, we
first investigated the binding of a more divergent but artificial A353R mutation under increasing salt concentrations. Surprisingly, NAP57 A353R retained SHQ1 markedly better on amylose resin than wild-type NAP57 (Fig. 4A), which was particularly conspicuous in the presence of 210 mM salt (compare lanes 5 and 6). Subsequent testing of the A353V patient mutation also showed it bound more SHQ1 than the wild-type, although not quite as noticeably (Fig. 4B).

However, quantification of the gel bands revealed a near two-fold (1.7 ± 0.4, n = 3) increase of the SHQ1/NAP57 ratio between wild-type and A353V mutant NAP57 (Fig. 4B). Therefore, this patient mutation enhances the affinity of NAP57 for its assembly factor SHQ1. Interestingly, testing of additional patient mutants revealed some with decreased affinity for SHQ1, whereas that to threonine (M350T) decreased it (Fig. 4C, compare lanes 5 and 6). In contrast, mutation of serine 121 to glycine barely impacted SHQ1 binding, consistent with the location of that serine in the center of the catalytic domain, which is not required for SHQ1 binding. Therefore, the S121G mutation may primarily affect the catalytic activity of NAP57 rather than the biogenesis of H/ACA RNPs. Regardless, our data demonstrate that DC mutations generally impact the interaction between NAP57 and the H/ACA RNP assembly factor SHQ1 in a hypo- or hypermorphic fashion marking X-linked DC as a potential RNP assembly deficiency.

**DISCUSSION**

X-linked DC is caused by mutations in NAP57, a core protein of all H/ACA RNPs. These mostly missense mutations concentrate in the same three-dimensional domain formed by the NT and CT of the protein. In a model of human H/ACA RNPs, the mutated amino acids do not compromise RNP...
integrity directly, as they are not contacting any other proteins of the particle nor the H/ACA RNA. In contrast, we document that the DC mutations impact the interaction of NAP57 with the H/ACA RNP assembly factor SHQ1.

To understand how both increased and decreased affinity of NAP57 for SHQ1 has the same dire consequences, it is important to consider the relationship and function of the two proteins.

SHQ1 functions in an obligate early step of H/ACA RNP biogenesis (13). Specifically, SHQ1 binds to newly synthesized NAP57, stabilizing it and preventing it from misfolding and degradation. In support of this notion, excess SHQ1 inhibits RNP assembly by sequestration of NAP57 and depletion of SHQ1 causes a loss of NAP57 due to degradation (13,20). Therefore, both increase and decrease of SHQ1 share the same consequence, reduced availability of NAP57 for RNP assembly. Apparently, SHQ1 and NAP57 live in a yin–yang relationship with specific cellular levels of each being crucial to cell viability. We now demonstrate that X-linked DC mutations cause an imbalance in the delicate relationship of NAP57 and SHQ1 by modulating their affinity for each other and thereby impairing H/ACA RNP biogenesis.

Moreover, a mouse model of X-linked DC is hypomorphic for NAP57 (23) and reduced expression of the yeast NAP57 ortholog, Cbf5p, suppresses tRNA-mediated gene silencing (24). Here we find that DC mutations in NAP57, whether hyper- or hypomorphic, reduce its availability for RNP assembly. In summary, our data show that through physical interaction SHQ1 is an important regulator of cellular levels of free NAP57 and that modulation of this interaction forms a potential molecular basis for X-linked DC. Having pinpointed this interaction surface as the culprit in X-linked DC makes it amenable to small molecule screens for future therapeutic approaches.

MATERIALS AND METHODS

Modeling of human H/ACA core proteins

The human GAR1–NAP57–NOP10 complex was modeled using the archaeal 2apo (25) and 2ey4 (12) structures and NHP2 using the structure of the related 15.5K protein 1e7k (14) as templates following our previously described protocol (26). The structural model of the entire H/ACA core protein complex was obtained using the following three steps. First, NHP2 was blindly docked onto the GAR1–NAP57–NOP10 complex (i.e. non-biased docking) using FT-dock 2.0 (27). The search yielded a sampling space composed of more than 10 000 docking poses that exhibited the best surface complementarity between GAR1–NAP57–NOP10 and NHP2. The search yielded a sampling space composed of more than 10 000 docking poses that exhibited the best surface complementarity between GAR1–NAP57–NOP10 and NHP2. Second, more than 250 docking poses located in a region close to NOP10 (i.e. NHP2 docking poses whose center of
mass was contained within a distance of 10 Å to any residue of NOP10) were selected, because we previously showed NOP10 to be required for NHP2 binding (9). Thirdly, because 1e7k (14), the template used to model NHP2, was co-crystallized with RNA, we requested that the equivalent structural region in NHP2 be exposed and oriented consistently to the RNA binding region of the NOP10–NAP57–GAR1 complex, reducing the number of docking poses to 45. The remaining docking poses were rescored by using statistical potentials (28) and further refined using MultiDock (29). The top 20 docking poses were rescored by using statistical potentials cing the number of docking poses to 45. The remaining

The viability of NAP57Δcat was confirmed by computational means as follows. Two structural models of NAP57Δcat were built using NAP57 as template by linking G90 to P288 either directly (G90-P288) or by an intermediate glycine (G90-G-P288). To assess the structural integrity of the models, a short molecular dynamic simulation was performed. The models were minimized and simulated during 2 ns in implicit solvent (GBSA) (30) using AMBER ff99 force-field (31) in the TINKER molecular modeling package (http://dasher.wustl.edu/tinker). The structural models remained stable throughout the simulation, i.e. all intermediate conformational snapshots dumped during the simulation had a root mean square deviation of less than 2.5 Å (Cα) upon structural superposition of the initial models (i.e. prior to simulation).

The quality of the models was evaluated using ProSa (32). Both models, G90-P288 and G90-G-P288, had comparable ProSa Z-scores, −4.24 and −4.45, respectively. Lower ProSa Z-scores indicate more stable structures and both scores fell within the typical range of native proteins of similar size. The models in Figure 1 were generated using MacPyMOL (www.pymol.org).

DNA constructs and expression, and protein–protein interaction assays

Most constructs and methods have been described previously (13). For bacterial expression, the following MBP-fusions were constructed: MBP-NT (pPG65, first 90 amino acids of NAP57), MBP-cat (pPG48, amino acids 91–287 corresponding to catalytic domain of NAP57), MBP-CT (pPG33, amino acids 289–514 constituting the CT of NAP57) and MBP-NAP57Δcat (pPG32, NAP57 with its catalytic domain excised, pSR55, fusion of amino acids 90–287). Unlike the full-length MBP-NAP57, pPG17 (13), these constructs lacked a CT hexa-histidine tag. All NAP57 point mutants were identical to full-length MBP-NAP57 (with a CT hexa-histidine tag) and generated by site-directed mutagenesis directly in pPG17, T49M (pPG59), T66A (pPG63), M350I (pPG61), M350T (pPG60) or in a different vector first, S121G (pPG62, pSR44), A353R (pPG53, pSR47) and A353V (pPG51, pSR45). Codon replacement of all NAP57 mutants was verified by sequencing. For transient expression in tissue culture cells, some of the NAP57 constructs above were fused to the lac-repressor (LacI) and to monomeric GFP: LacI-cat (pPG49), LacI-CT (pPG30), LacI-NAP57Δcat (pPG23) and GFP-NAP57Δcat (pSR529).

Transient transfection, bacterial protein expression and purification, and protein–protein interactions in vitro (amylose resin pull-downs) and in vivo (nuclear tethering assay) were as we described previously (13,16). Gel bands were quantified using NIH ImageJ software.

ACKNOWLEDGEMENTS

We are grateful to Sujayita Roy and Sudhamani Ramamurthy for construction of some of the plasmids and to Susan Smith for critical comments on the manuscript.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by a grant from the National Institutes of Health (HL079566) to U.T.M.

REFERENCES


