Neuroscience
Physics of Living Systems

C9orf72 polyPR directly binds to various nuclear transport components

Hamidreza Jafarinia, Erik Van der Giessen, Patrick R. Onck

Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747AG Groningen, The Netherlands

Reviewed Preprint

Revised by authors after peer review.

About eLife's process

Reviewed preprint version 2
December 27, 2023 (this version)

Reviewed preprint version 1
September 12, 2023

Posted to preprint server
June 22, 2023

Sent for peer review
June 7, 2023

Abstract

The disruption of nucleocytoplasmic transport (NCT) is an important mechanism in neurodegenerative diseases. In the case of C9orf72-ALS, trafficking of macromolecules through the nuclear pore complex (NPC) might get frustrated by the binding of C9orf72-translated arginine-containing dipeptide repeat proteins (R-DPRs) to the Kapβ family of nuclear transport receptors. Beside Kapβs, several other types of transport components have been linked to NCT impairments in R-DPRs expressed cells, but the molecular origin of these observations has not been clarified. Here, we adopt a coarse-grained molecular dynamics model at amino-acid resolution to study the direct interaction between polyPR, the most toxic DPR, and various nuclear transport components to elucidate the binding mechanisms and provide a complete picture of potential polyPR-mediated NCT defects. We found polyPR to directly bind to several isoforms of the Impα family, CAS (the specific exporter of Impα) and RanGAP. We observe no binding between polyPR and Ran. Longer polyPRs at lower salt concentrations also make contact with RanGEF and NTF2. Analyzing the polyPR contact sites on the transport components reveals that polyPR potentially interferes with RanGTP/RanGDP binding, with nuclear localization signal (NLS)-containing cargoes (cargo-NLS) binding to Impα, with cargo-NLS release from Impα, and with Impα export from the nucleus. The abundance of polyPR binding sites on multiple transport components combined with the inherent polyPR length dependence makes direct polyPR interference of NCT a potential mechanistic pathway of C9orf72 toxicity.

eLife assessment

This study provides an **important** starting point for unraveling the molecular basis of the pathological phenotypes of the repeat expansion in the gene associated with open reading frame 72 in human chromosome 9. The coarse-grained simulation method used by the authors goes beyond the state of the art, investigating a **compelling** number of binding partners. The evidence supporting the claims of the authors is **solid**, although experimental validation of the results would strengthen the major conclusions of the work. The work will be of broad interest to biophysicists and biochemists.

Introduction

The C9orf72 G4C2 hexanucleotide repeat expansion is the most common genetic mutation in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) [1, 2]. This expansion can be translated into five types of dipeptide repeat proteins (DPRs): polyPR, polyGR, polyGA, polyGP, and polyPA [3]. The positively-charged arginine-containing DPRs (R-DPRs) show the highest levels of toxicity in different cell and animal models [4-10] with polyPR known to be the most toxic DPR [5, 9, 11]. R-DPRs have been linked to a wide variety of cellular defects [12-16], but a growing body of evidence suggests that neurodegenerative diseases, including C9orf72 ALS/FTD (C9-ALS/FTD), may be caused by disruption of nucleocytoplasmic transport (NCT) [17-20]. At the same time, many transport components that play a prominent role in NCT have been identified to function as modifiers of G4C2/DPR toxicity [5, 10, 21].

The regulated trafficking of proteins and RNA between the nucleus and cytoplasm occurs through nuclear pore complexes (NPCs) embedded in the nuclear membrane [22, 23]. The NPC is lined with intrinsically disordered phenylalanine-glycine-rich nucleoporins (FG-Nups) that collectively function as a selective permeability barrier. Small molecules rapidly diffuse through the NPC, but the passage of larger cargoes across the barrier needs to be facilitated by their binding to nuclear transport receptors (NTRs) [24-26]. The β-karyopherin (Kapβ) family is the largest class of NTRs and includes both import and export receptors [27]. Another essential regulator of NCT is the GTPase Ran, a small protein bound to guanosine triphosphate (GTP) in the nucleus and to guanosine diphosphate (GDP) in the cytoplasm [24]. The directionality of NCT is mediated by the RanGTP-RanGDP gradient over the nuclear envelope, which is preserved by the cytoplasmic GTPase-activating protein RanGAP and the nuclear guanine nucleotide exchange factor RanGEF [24, 28].

In the import cycle, importins bind their cargoes directly through a nuclear localization signal (NLS) encoded on a cargo. Importin β1 (Impβ1) can also recruit importin α (Impα) that functions as a cargo-adaptor protein. Impα binds to Impβ1 through its N-terminal importin β-binding (IBB) domain [29, 30]. The importin-NLS-cargo complex then shuttles to the nucleus. The binding of RanGTP to importin in the nucleus disassembles the importin-NLS-cargo complex and the RanGTP-importin complex is recycled to the cytoplasm. When cargo is bound to Impβ1 via Impα, RanGTP dissociates Impβ1 from the Impα-NLS-cargo. This triggers a competition between the flexible IBB domain of Impα and NLS-cargo for binding to Impα, thus facilitating the dissociation of the NLS-cargo. Nucleoporins such as Nup50/Nup2 also catalyze this process by binding to Impα and accelerating the dissociation rate of the cargo-NLS [22]. The specific receptor CAS bound to RanGTP is required to export Impα to the cytoplasm. It has been proposed that CAS first displaces Nup50/Nup2 from Impα after which the RanGTP-CAS-Impα returns to the cytoplasm. The hydrolysis of RanGTP to RanGDP in the cytoplasm by RanGAP disassembles the RanGTP-importin and RanGTP-CAS-Impα complexes [22]. RanGDP is transported back to the nucleus by nuclear transport factor 2 (NTF2) where the RanGDP-NTF2 complex dissociates when RanGEF regenerates RanGTP [22]. In the export cycle, RanGTP promotes the loading of cargoes with nuclear export signal (NES) to the exportin in the nucleus. The resulting RanGTP-exportin-NES-cargo complex moves to the cytoplasm. Once there the complex is disassembled by RanGAP which hydrolyses RanGTP to RanGDP [24].

In a recent study we have analyzed the binding of polyPR to the Kapβ family of importins and exportins [13, 16, 31] by using coarse-grained (CG) molecular dynamics simulations [32]. Depending on its length, polyPR can interact with several cargo-, IBB-, RanGTP-, and FG-Nup-binding sites on the Kapβs [32]. Beside Kapβs, there is evidence for direct binding of Impα isomers with R-DPRs [16]. Some regulators of the Ran cycle are also affected in R-DPR-mediated toxicity, where R-DPRs have been shown to cause mislocalization and abnormal accumulation of

RanGAP [33], and mislocalization of Ran and RanGEF in cell culture models [5, 33, 34]. RanGAP and RanGEF also appear to be modifiers of R-DPR toxicity in genetic studies [5, 10, 11]. It is not clear, however, whether these effects arise from a direct interaction of R-DPRs with NCT components. The aim of the current work is to extend the findings of [32] by investigating the interaction between polyPR and various NCT components by means of CG molecular dynamics computations.

Results and discussion

Coarse-grained models of nucleocytoplasmic transport components

We use the residue-scale CG molecular dynamics approach developed and applied earlier to study DPR phase separation [35] and the direct binding of polyPR to numerous members of the Kapβ family [32]. In the present work we investigate the interaction of polyPR with unbound human Impα isomers (Impα1, Impα3, Impα5, Impα7), Ran, CAS (the specific exporter of Impα), RanGEF, and NTF2. We also include KAP60 (homolog of Impα), Cse1 (homolog of CAS), and RanGAP from yeast, since contrary to the human homologs, the crystal structures of the KAP60-Cse1 complex and the RanGAP-RanGppNHp complex (fission yeast RanGAP bound to the non-hydrolyzable form of human RanGTP) are available in the protein data bank. This enables us to investigate a possible polyPR interference with Impα export and the RanGAP function in the model system of yeast. Moreover, yeast has been employed previously to study NCT defects caused by DPRs [5, 36]. More details about the selected transport components can be found in table S2 of the SI.

In our one-bead-per-amino-acid (1BPA) CG models of transport components, each residue is represented by a single bead at the position of the alpha-carbon atom. The overall tertiary structure of the NCT components is preserved through a network of stiff harmonic bonds, and the distribution of charged and aromatic residues are included in the model. The 1BPA force field correlates with experimental findings for polyPR-Kapβs interactions [32]. The CG models of the NCT components studied are shown in **figure 1**. The Impα isomers contain a flexible N-terminal IBB domain, followed by a helical core that is constructed from 10 Armadillo (ARM) repeats each consisting of three alpha helices. The NLS binding sites are located on the concave surface of the helical core [37]. The IBB domain has an autoinhibitory role and when it is not bound to Impβ1, it binds to the ARM concave core and competes with NLS binding [22]. To simplify the study of possible binding between polyPR and NLS binding sites of Impα, the CG models of Impα isomers are built without their N-terminal IBB domains and referred to as ImpαΔN. This approach enables easier investigation of polyPR's interaction with NLS binding sites. Similar to the importins and exportins, CAS and Cse1 are constructed from HEAT repeats, with each repeat consisting of two antiparallel α-helices, named A and B, connected by linkers of different lengths [38]. RanGEF is mainly constructed from β-strands and has an overall appearance of a seven-bladed propeller with each blade consisting of a four-stranded antiparallel β-sheet [39]. The RanGAP model of fission yeast used in our study is constructed from eleven leucine-rich repeats (LRRs) forming a symmetric crescent followed by a highly negatively-charged C-terminal region. Each LRR motif consists of a β strand-α helix hairpin unit [40]. NTF2 is a homodimer with each monomer consisting of a β-sheet and three α-helices [41]. The CG model of Ran is based on the nucleotide-free state of the molecule. More details about the CG models and forcefield are provided in the Methods section and section 1 of the SI.

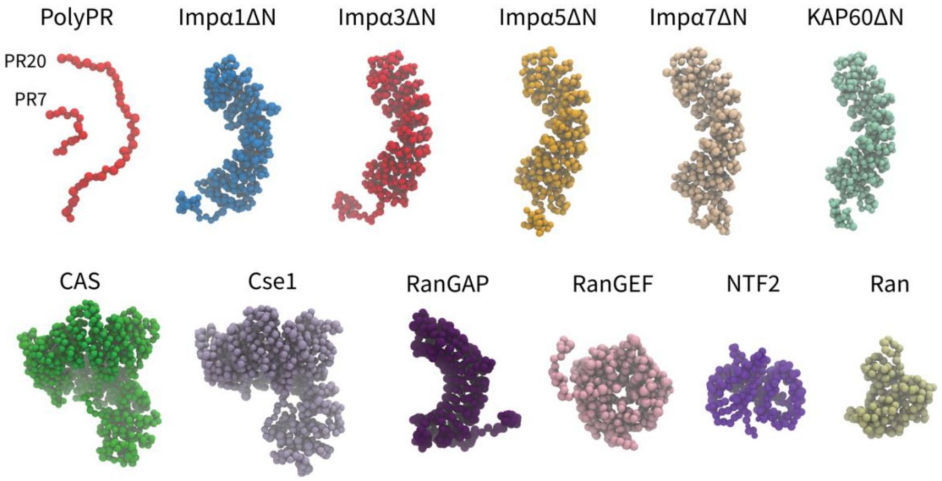

Figure 1

Coarse-grained models of polyPR and transport components used in this study

One-bead-per-amino acid (1BPA) representation of PR7, 20 and the various transport components modeled in the current study. These are several members of the Impα family (excluding the N-terminal IBB domain), specific exporters of Impα (CAS and Cse1), RanGAP, RanGEF, NTF2, and Ran. Details regarding the CG models and protein sequences are listed in tables S2 and S3.

PolyPR binds to several transport components through electrostatic interactions

In investigating the direct interaction with the transport components shown in **figure 1** we include the potential effect of polyPR length, as various studies have found that the repeat length of DPRs strongly correlates with toxicity [4, 42-44]. Simulations of polyPR with varying numbers of repeat units, i.e., PR7, PR20, and PR50, are performed at two salt concentrations: 200 mM, similar to previous in vitro experiments performed for Kapβ importins and Impα isomers interacting with R-DPRs [16], and a lower ion concentration of 100 mM to study the effect of salt concentration.

To quantify the interaction between polyPR and the transport components, we calculate the time-averaged number of contacts $C_t$using a cutoff of 1 nm. The number of contacts is normalized by the sequence length of the transport components (N_{TC}) and by the polyPR length (N_{PR}). The normalized number of contacts is plotted against the charge parameter

$$NCPR - fM/R_{\mathrm{g}}, \qquad (1)$$

where the net charge per residue *NCPR* is the total charge of the transport component (in units of elementary charge e) divided by its sequence length, M (in units of e · nm) is the time-averaged total dipole moment, and R_g (in units of nm) is the time-averaged radius of gyration of the transport components in isolation. The dimensionless parameter f is a free parameter which is calculated to be 0.0036 for the best linear fit in **figures 2a** for PR20 and PR50, based on the quality of fit (R^2), see figure S1. The linear correlation observed in **figure 2a** confirms an electrostatically driven interaction between polyPR and the transport components. Moreover, it highlights the importance of the spatial distribution of charge over the transport components, as characterized here through the dipole moment. CAS and Cse1, the specific exporters of Impα in human and yeast, respectively, are constructed from HEAT repeats and have a super-helical conformation similar to the Kapβs studied before [32]. We therefore present the results for these two cases jointly with the results for the Kapβ set (taken from [32]) in **figure 2b**. For this set, the best fit is obtained for $f = 0$ (see figure S1), which indicates the dominant role of *NCPR* for the number of contacts between polyPR and the Kapβs, CAS and Cse1. We attribute this behavior to the structural characteristics of Kapβs, particularly the superhelical structure which features inner and outer surfaces with differing charge distributions. Importantly, this structural arrangement creates an inner surface characterized by a strong negative electrostatic potential. As demonstrated in our previous work, polyPR predominantly binds to this negatively charged cavity within Kapβs. Consequently, the separation of charges on the Kapβ surface becomes less influential compared to the overall charge.

As can be seen in **figure 2a**, RanGAP features a much larger number of contacts with polyPR than the other transport components, which can be related to the higher negative net charge and the higher dipole moment of this molecule (see the dipole moments *M* and *NCPR* of the transport components in figures S2 and S3 of the SI). In contrast, polyPR makes a negligible number of contacts with Ran since it has a relatively low dipole moment and no net charge. For NTF2 the value of *NCPR* is − 0.047 e, comparable to several members of the ImpαΔN family, but the lower dipole moment of NTF2 results in a lower number of contacts with polyPR. The binding of PR20 to the ImpαΔN isomers is consistent with the Impα1 and Impα3 binding to R-DPRs found in experiment [16], despite the fact that the N-terminal IBB domains of Impα1,3 are excluded from our CG models. At 200 mM salt concentration, polyPR does not bind to NTF2 and Ran. PolyPR contact with RanGEF is also very low at this salt concentration. Reducing the salt concentration to 100 mM, increases the number of contacts, clearly indicating that electrostatic force is the main driver for binding. At this lower salt concentration, PR7, PR20, and PR50 make contact with RanGEF, but contact with NTF2 is only observed for longer polyPR chains.

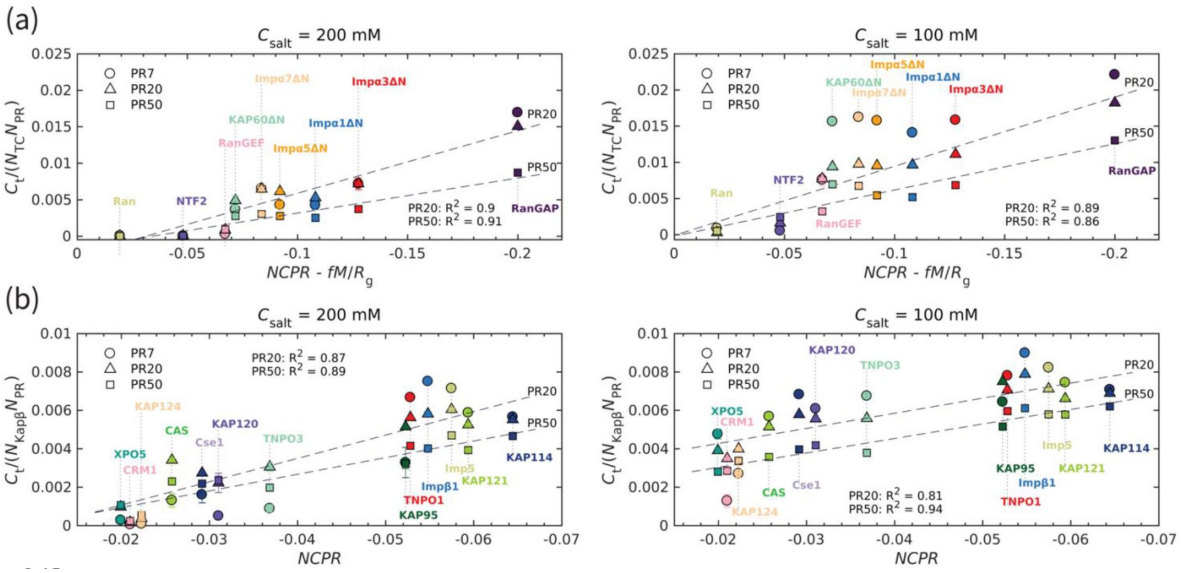

Figure 2

The transport component,s net charge per residue and dipole moment, together with polyPR length, affect polyPR interaction with various nuclear transport components

(a) and (b) show the normalized time-averaged number of contacts C_t for the interaction between polyPR with 7, 20, and 50 repeat units with different types of transport components. The results are shown for monovalent salt concentrations of C_{salt} = 200 mM (left panels) and C_{salt} = 100 mM (right panels). Subfigure (a) shows the results for the transport components shown in **figure 1**, excluding the specific exporters of Impα: CAS and Cse1. A linear correlation is observed between the normalized C_t and $NCPR - fM/R_g$ with f calculated to be 0.0036 for the best fit. The net charge per residue $NCPR$ is in units of elementary charge e, the dipole moment M is in units of e · nm, and the radius of gyration R_g is in units of nm. Subfigure (b) shows the results for the Kapβ data set (data points taken from [32]) together with CAS and Cse1. For this case, a linear correlation between the normalized C_t and $NCPR$ is observed. The dashed lines show linear fits for PR20 and PR50, see table S4 for the linear equations of the fits. The fits for PR7 resulted in R2 values of 0.89 (a) and 0.83 (b) for 200M and of 0.7 (a) and 0.59 (b) for 100 mM. Because of the low R2 values for 100 mM, the fits for PR7 are not shown. The error bars denote the standard error of the mean. Where error bars are invisible, they are smaller than the marker size.

For transport components with higher absolute values of *NCPR* and *NCPR* − fM/R_g, the number of contacts increases with increasing polyPR length. However, as can be seen in **figure 2**, the number of contacts per unit length of polyPR is often seen to be lower for longer polyPRs especially for the lower salt concentrations where polyPR strongly binds to certain transport components, see e.g. the results in **figure 2a** for the ImpαΔN family and RanGAP at 100 mM salt concentration. This is due to the fact that most of the residues make contact with the target protein for shorter polyPRs, while for longer polyPRs only some parts of the chain are in contact with the transport components and other regions make less or no contact.

PolyPR interacts with important binding sites of transport components

In order to gain a better understanding of how polyPR interacts with each transport component, we examined the polyPR contact probability of each residue in the sequence of the transport component. The contact probability of each residue is defined as the probability of having at least one polyPR residue within its 1 nm proximity. **Figure 3a** reveals our results of the interaction between PR7 and PR50 with Impα1ΔN, KAP60ΔN, Cse1, RanGAP, RanGEF and NTF2 (for the other transport components, see figure S4). These findings indicate that a longer polyPR makes contact with a larger number of sites, and also exhibits a higher contact probability with individual residues compared to a shorter polyPR. We also observe that some regions at the C-terminal ends of the ImpαΔN family and RanGAP are permanently bound to polyPR. We also compare the polyPR binding sites with known binding sites of transport components (according to the PiSITE webserver [45]), highlighting them at the bottom of each subfigure. **Figure 3b** displays the number of contact residues shared between polyPR and the native binding partners of each transport component. As expected, the general trend is that the number of shared binding sites increases with increasing polyPR length. PolyPR interacts with the ImpαΔN family at several known cargo-NLS, Nup50/Nup2, and Cse1 binding sites. Longer polyPRs exhibit a significantly stronger interaction with cargo-NLS binding sites of ImpαΔN compared to shorter polyPRs, which only interact with a limited number of sites. However, we observe that both short and long polyPRs bind to Nup50/Nup2, Cse1, and RanGTP binding sites particularly those located near the C-terminal end of ImpαΔN. In the case of Cse1, we observe polyPR binding to Impα and RanGTP binding sites. PolyPR interacts with the known RanGTP binding sites of RanGAP. We also show that polyPR is able to bind to the highly negatively-charged region in the C-terminal domain of RanGAP that follows the LRR domain. It has been suggested that this negatively-charged region is in close proximity to a positively-charged region in Ran (in the complex formed by Ran and RanGAP) and plays a role in RanGTP hydrolysis [46, 47]. Unfortunately, there is no crystal structure for this region in the PDB structure and thus the binding sites are not known. In the case of RanGEF, a longer polyPR interacts with a high percentage of the known Ran binding sites. For NTF2 we observe polyPR interaction with RanGDP binding sites. PolyPR makes negligible contacts with nucleotide-free state of Ran. For CAS the binding sites are not known. Therefore, these two cases are excluded from **figure 3b**.

The findings presented in **figure 3b** and S4, along with previous research on the interaction between polyPR and Kapβs (importins and exportins) [32], lead to the following mechanistic understanding for the direct effect of polyPR on NCT as illustrated in **figure 4**. In this figure the native binding interactions that are affected by polyPR are indicated by red arrows and those that are unaffected by grey arrows. In the import cycle, NLS-cargoes bind to Kapβs directly or indirectly through adaptor proteins such as Impα isomers. PolyPR may impede the loading of cargo to Kapβs and Impα isomers (as shown in inset A) by binding to the cargo-NLS sites. For Impα isomers, the binding of polyPR to cargo-NLS binding sites is mostly limited to longer chains, see **figure 3b**. RanGTP disassembles the import complex (cargo-Kapβ or cargo-Impα-Kapβ), and Nup2/Nup50 facilitate the disassembly of the cargo-Impα complex in the nucleus. PolyPR binding to the RanGTP binding sites on the Kapβs and the Nup2/Nup50 binding sites on Impα, see **figure 3b**, could result in defects in the dissociation of cargo/cargo-Impα from Kapβ and cargo from

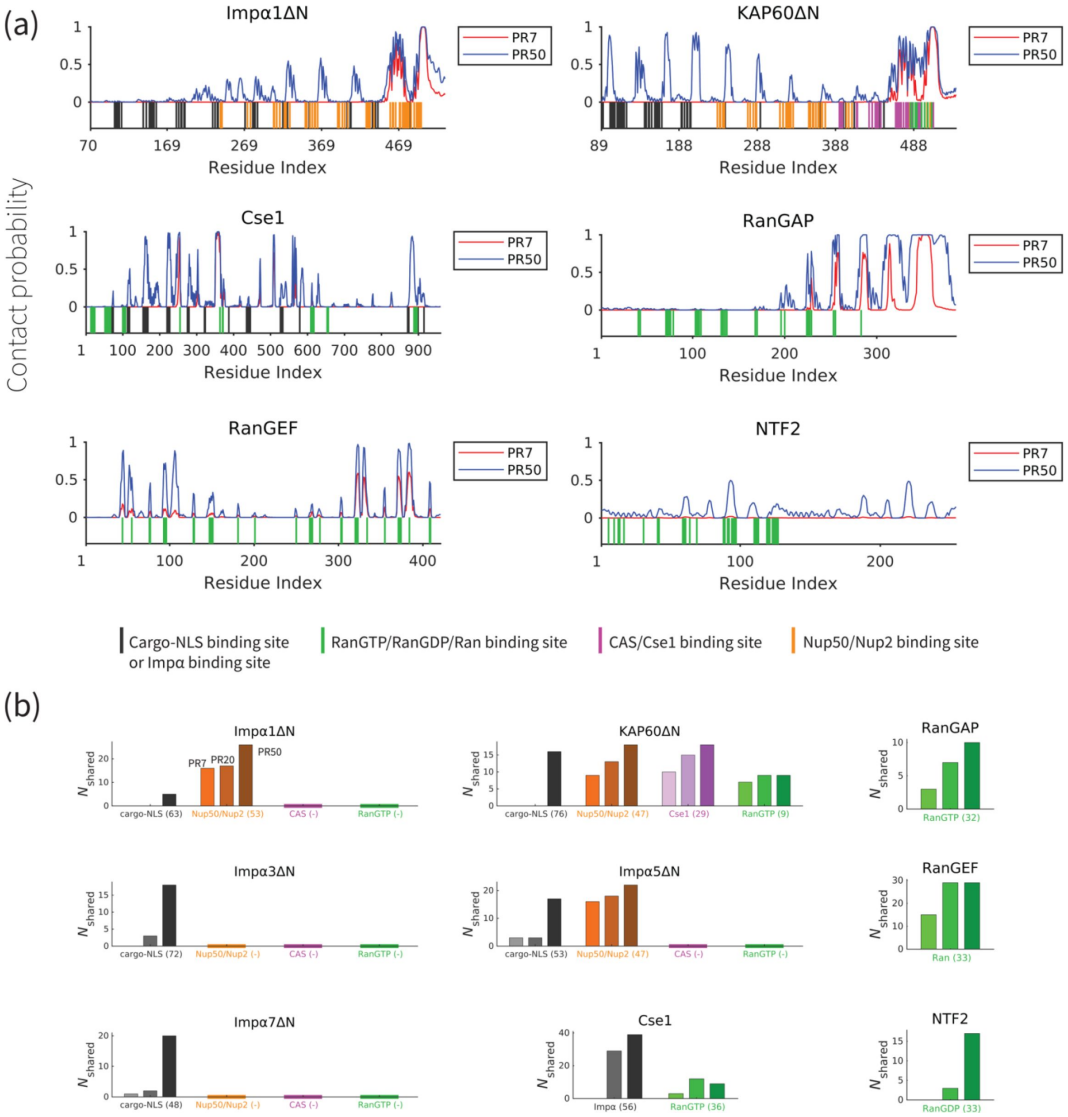

Figure 3

PolyPR interacts with several known binding sites of nuclear transport components in a length-dependent manner

(a) The contact probability for each residue in the sequence of transport components interacting with polyPR. The plot displays the contact probability for six transport components: Impα1ΔN, KAP60ΔN, Cse1, RanGAP, RanGEF and NTF2 at a salt concentration of 100 mM. Results for Impα3ΔN, Impα5ΔN, Impα7ΔN, CAS, and Ran are shown in figure S4. Each figure shows two curves for PR7 and PR50. The bottom part of each figure shows the binding sites for NLS-cargo, Impα, CAS/Cse1, RanGTP, and Nup50/Nup2 using different colors. These binding sites are obtained from the crystal structures of the bound states of transport components in the Protein Data Bank using PiSITE (see table S2 of the SI for more details). For each transport component the following binding sites are marked. For the Impα family: NLS-cargo (vertical black lines), CAS/Cse1 (vertical purple lines) and Nup50/Nup2 (vertical orange lines) binding sites. For CAS/Cse1: Impα (vertical black lines) and RanGTP (vertical green lines) binding sites. For RanGAP: RanGTP (vertical green lines) binding sites. For RanGEF: Ran (vertical green lines) binding sites. For NTF2: RanGDP (vertical green lines) binding sites. The Ran binding sites marked for RanGEF are taken from the RanGEF-Ran complex (an intermediate step in the RanGEF function).
(b) The number of shared contact sites between polyPR and the binding partners of the transport components, referred to as N_{shared}, are plotted for PR7, PR20, and PR50. In each bar plot, the numbers inside the parentheses on the horizontal axis shows the number of known binding sites obtained from PiSITE. If there is no known binding site, a (-) mark is used instead. The results for PR7, PR20, and PR50 are reported from left to right for each set of bar plots. The bars with darker colors represent longer polyPR chains.

Impα (as shown in insets B and C). CAS/Cse1 export Impα by forming a complex with Impα and RanGTP. Findings in **figure 3b** show that polyPR also interacts with certain CAS/Cse1 binding sites that recognize Impα and RanGTP, possibly interfering with the formation of the RanGTP-CAS/Cse1-Impα complex (as shown in inset D). RanGAP plays a crucial role in the NCT cycle by mediating the hydrolysis of RanGTP to RanGDP, leading to the disassembly of export complexes (RanGTP-importin, RanGTP-CAS/Cse1-Impα, and RanGTP-exportin-cargo). The relatively high number of contacts between polyPR and RanGAP (see **figure 2a**), as well as polyPR binding to RanGTP binding sites on RanGAP (see **figure 3b**), suggest a possible defect in the dissociation of the RanGTP-importin and RanGTP-CAS/Cse1-Impα complexes in the import and Ran cycle, and of the RanGTP-exportin-cargo complex in the export cycle (as shown in inset E). Following the hydrolysis of RanGTP to RanGDP in the cytoplasm, RanGDP is transported back to the nucleus by nuclear transport factor 2 (NTF2). Once in the nucleus, the RanGDP-NTF2 complex dissociates when RanGEF exchanges GDP for GTP in Ran. **Figure 3b** shows that longer polyPRs interact with the Ran binding sites of RanGEF and NTF2. We therefore suggest that longer polyPRs may also interfere with the Ran cycle by hindering the loading of RanGDP to NTF2 and the exchange of GDP to GTP in Ran by RanGEF (as shown in insets G and F). The interaction between polyPR and NTF2 is distinguished by a relatively lower number of contacts (**figure 2a**) and contact probabilities for individual NTF2 residues (**figure 3b**). These findings suggest a lower likelihood of polyPR interfering with NTF2 function compared to the other functions outlined in **figure 4**. In the export cycle, as examined previously [32], polyPR binding to RanGTP and FG-Nup binding sites may affect cargo-loading onto exportins (as shown in inset H) and the transport of exportins through the NPC (as shown in inset I). It should be noted that the mechanisms proposed in **figure 4** do not hold equal weight, and the relative contributions based on the number of contacts (presented in **figure 2**), and the number of contacts with important binding sites (presented in **figure 3b** and S4) should be considered when comparing the relative significance of the suggested mechanisms.

Conclusion

In this study we used coarse-grained molecular dynamics simulations to show that polyPR binds to several nuclear transport components. Similar to the interaction of polyPR to the Kapβ family, the interaction of polyPR to other transport components is driven by electrostatic interactions and depends sensitively on polyPR length. Reducing the salt concentration or increasing the polyPR length increases the number of contacts with different transport components. The effect of polyPR length suggests a molecular basis for the more toxic nature of longer polyPRs in animal and cell models [4, 42-44].

We observed polyPR binding to several members of the Impα family, CAS, Cse1, and RanGAP yet no binding to Ran. PolyPR strongly binds to a highly negatively charged region in the C-terminal domain of fission yeast RanGAP. The human RanGAP also contains a similar highly negatively-charged region. Our simulations therefore suggest that a direct interaction may contribute to the observed polyPR-mediated accumulation and mislocalization of RanGAP in HeLa cells [33]. PolyPR also binds to RanGEF and NTF2 at lower salt concentrations or when the polyPR length is large enough. We also showed that incorporating the dipole moment leads to an improved fit for the transport components analyzed in **figure 2a**, suggesting that binding is influenced not only by the net charge per residue (as previously observed for Kapβs [32]) but also by the spatial separation of charges on the transport component.

We showed that polyPR interacts with important binding sites of different transport components in a polyPR-length-dependent manner, with polyPR interaction with RanGTP/RanGDP binding sites being a common feature between the transport components. This suggests a strong polyPR interference with the Ran gradient across the nuclear envelop. For the ImpαΔN family, we observe polyPR interaction with cargo-NLS and Nup2/Nup50 binding sites. In the case of KAP60ΔN (yeast

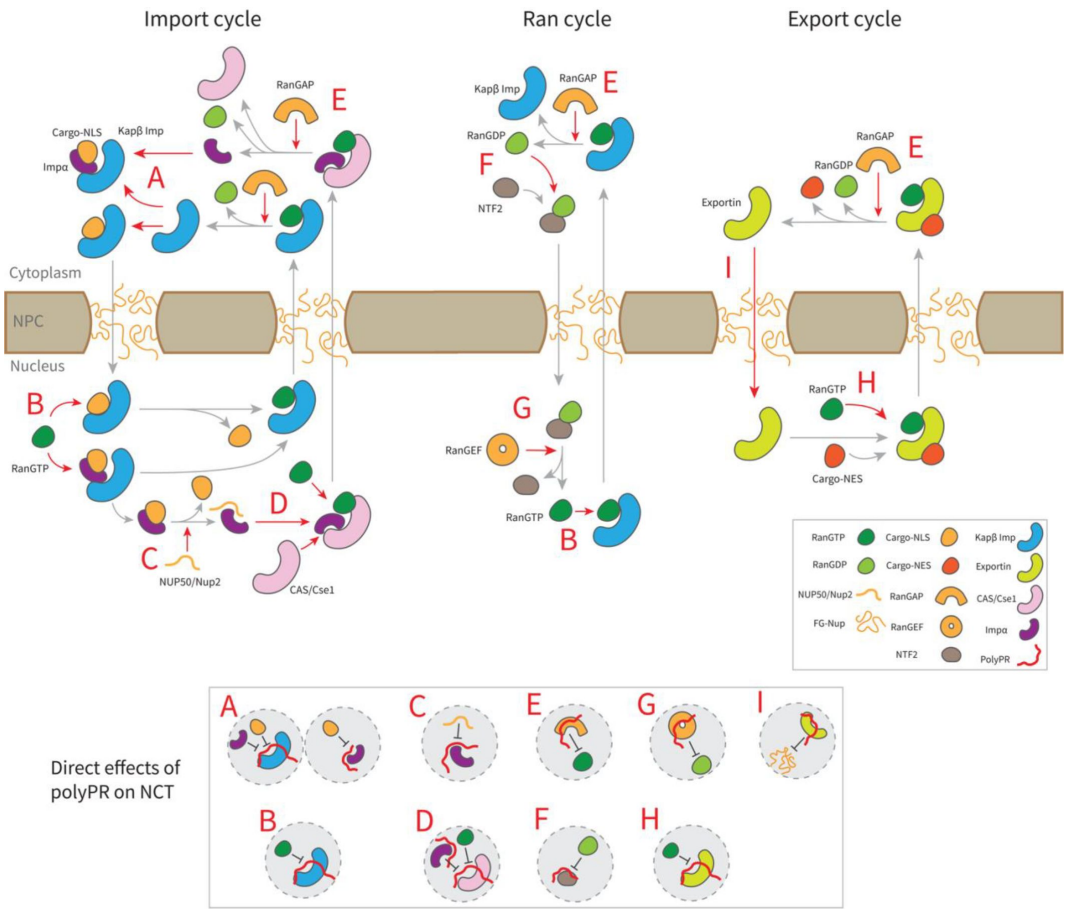

Figure 4

Suggested molecular mechanism of polyPR interference with the native function of transport components in the nucleocytoplasmic transport cycle.

(a) Proposed mechanistic pathways of polyPR interference with the import cycle (left panel), Ran cycle (middle panel), and export cycle (right panel). Steps in the NCT cycle are represented with grey arrows, and a red dashed arrow indicates where polyPR may interfere with the transport cycle. The letters A-H are used to illustrate how polyPR may disrupt the function of the transport components. Each letter corresponds to a mechanistic mechanism shown at the bottom of the figure in grey circles. It should be noted that the proposed mechanisms are not equally significant. The relative significance of the suggested molecular mechanisms can be obtained by considering their relative contributions based on the number of contacts and the number of contacts with important binding sites as presented in **figures 2** and **3**, respectively.

homolog of Impα), we observe polyPR interaction with Cse1 binding sites. For Cse1 (yeast CAS), we also observed polyPR interacting with Impα binding sites. These findings suggest polyPR interference with the cargo-NLS association and disassociation with Impα, and the export of Impα.

In conclusion, we showed a pronounced direct binding interaction between polyPR and a surprisingly large number of transport components. By integrating our findings with previously reported data, this work proposes a molecular model that explains how the binding of polyPR might interfere with distinct stages of the transport cycle. The intrinsic length dependence of polyPR binding to important binding sites of many transport components promotes this mechanism to a potential target for therapeutic interventions. Overall, our results offer a basis for future research that aims to explore the impact of C9orf72 R-DPRs on NCT disruption and the subsequent downstream consequences.

Methods

Coarse-grained model

We adopt a one-bead-per-amino acid (1BPA) force field to study polyPR interaction with nucleocytoplasmic transport components. This 1BPA approach has been initially developed to simulate disordered FG-Nups [25, 48-53], and extended later to study the phase separation of DPRs [35], and the interaction of polyPR with Kapβs [32]. The force field potentials and parameters in this study are identical to those employed in [32]. The bonded potentials for polyPR are residue and sequence specific. For non-bonded polyPR-polyPR interactions, the force field accounts for hydrophobic/hydrophilic and electrostatic interactions. The crystal structure of the transport components is maintained using a stiff harmonic potential $\varphi_{\text{network}} = K(r - b)^2$, where K is 8000 kJ/mol/nm^2 and b is the distance between the amino acid beads in the crystal structure. A bond is made between the beads if b is less than 1.4 nm. The unresolved regions in the crystal structure from the Protein Data Bank, and the regions with a lower prediction score (< 70 pLDDT) from Alpha Fold [54, 55] are included in the CG model as disordered regions. The CG model of fission yeast RanGAP includes two alpha helices in the C-terminal domain, as predicted by AlphaFold. The polyPR interactions with transport components are classified into three categories: (1) electrostatic interactions, (2) cation-pi interactions, and (3) excluded volume interactions. Our force field also accounts for the screening effect of ions. More details about the CG models and forcefield are provided in table S2 and section 1 of the SI.

Simulation and analysis

Langevin dynamics simulations are performed at 300 K at monovalent salt concentrations of 100 mM and 200 mM in NVT ensembles with a time-step of 0.02 ps and a Langevin friction coefficient of 0.02 ps^{-1} using GROMACS version 2018. Simulations are performed for at least 2.5 μs in cubic periodic boxes, and the last 2 μs are used for analyzing the interaction between polyPR and the transport components. The error bars in **figure 2** are standard errors of the mean (SEM) calculated from block averaging with three blocks at equilibrium. The binding sites are obtained from the crystal structures of the bound states of transport components in the Protein Data Bank using PiSITE [45]. This web-based database provides interaction sites of a protein from multiple PDBs including similar proteins. The RanGTP binding data (vertical green lines) in **figure 3** and S4 contains binding residues for both RanGTP and RanGppNHp, the non-hydrolysable form of RanGTP. The time-averaged number of contacts between the polyPR and transport components in **figure 2** is obtained by summing the number of contacts per time frame (i.e. the number of polyPR/transport components residue pairs that are within 1 nm) over all frames and dividing by the total number of frames. The contact probability for each transport component residue is the probability of having at least one polyPR residue within 1 nm proximity of the transport component residue. The contact probability is calculated for each transport component residue by dividing the number of frames for which this contact criterion is satisfied, by the total number of

frames. In figure S5 we show that conducting longer simulations does not significantly affect the contact probabilities presented in **figure 3a** (data shown for PR50 binding to Impα1ΔN and RanGAP), confirming convergence of our computations. Residue *i* is considered to be a contact site if the contact probability for this residue is larger than 0.10. N_{shared} is the number of transport component residues that make contact with polyPR (obtained in our simulations) and at the same time are known for recognition of native binding partners (according to PiSITE). The time-averaged total dipole moment of the CG models of transport components is calculated using gmx dipole in GROMACS.

Data availability

The data and scripts of this study are available from the corresponding author upon reasonable request.

Author contributions

H.J., E.Vd.G., and P.R.O. designed research; H.J. performed and analyzed research; and H.J., E.Vd.G, and P.R.O. wrote the paper.

References

1. Renton A.E., et al. (2011) **A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD** *Neuron* **72**:257–268

2. DeJesus-Hernandez M., et al. (2011) **Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS** *Neuron* **72**:245–256

3. Mori K., et al. (2013) **The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS** *Science* :1335–1338

4. Mizielinska S., et al. (2014) **C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins** *Science* **345**:1192–1194

5. Jovičič A., et al. (2015) **Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS** *Nat. Neurosci* **18**:1226–1229

6. Kwon I., et al. (2014) **Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells** *Science* **345**:1139–1145

7. Zhang Y.J., et al. (2019) **Heterochromatin anomalies and double-stranded RNA accumulation underlie C9orf72 poly(PR) toxicity** *Science* **363**

8. Zhang Y.J., et al. (2018) **Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis** *Nat. Med* **24**:1136–1142

9. Wen X., et al. (2014) **Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate invitro and invivo neuronal death** *Neuron* **84**:1213–1225

10. Boeynaems S., et al. (2016) **Drosophila screen connects nuclear transport genes to DPR pathology in c9ALS/FTD** *Sci. Rep* **6**:20877–20877

11. Lee K.H., et al. (2016) **C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles** *Cell* **167**:774–788

12. Shi K.Y., et al. (2017) **Toxic PRn poly-dipeptides encoded by the C9orf72 repeat expansion block nuclear import and export** *Proceedings of the National Academy of Sciences of the United States of America* **114**:201620293–E1117

13. Hayes L.R., Duan L., Bowen K., Kalab P., Rothstein J.D. (2020) **C9orf72 arginine-rich dipeptide repeat proteins disrupt karyopherin-mediated nuclear import** *Elife* **9**

14. Balendra R., Isaacs A.M. (2018) **C9orf72-mediated ALS and FTD: multiple pathways to disease** *Nature Reviews Neurology* **14**:544–558

15. Freibaum B.D., et al. (2015) **GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport** *Nature* **525**:129–133

16. Hutten S., et al. (2020) **Nuclear Import Receptors Directly Bind to Arginine-Rich Dipeptide Repeat Proteins and Suppress Their Pathological Interactions** *Cell Rep* **33**

17. Kim H.J., Taylor J.P. (2017) **Lost in Transportation: Nucleocytoplasmic Transport Defects in ALS and Other Neurodegenerative Diseases** *Neuron* **96**:285–297

18. Mihevc S.P., et al. (2016) **Nuclear trafficking in amyotrophic lateral sclerosis and frontotemporal lobar degeneration** *Brain : a journal of neurology* **140**:13–26

19. Boeynaems S., Bogaert E., Van Damme P., Van Den Bosch L. (2016) **Inside out: the role of nucleocytoplasmic transport in ALS and FTLD** *Acta Neuropathol* **132**:159–173

20. Jovičić A., Paul J.W., Gitler A.D. (2016) **Nuclear transport dysfunction: a common theme in amyotrophic lateral sclerosis and frontotemporal dementia** *Journal of neurochemistry* **138**:134–144

21. Kramer N.J., et al. (2018) **CRISPR-Cas9 screens in human cells and primary neurons identify modifiers of C9ORF72 dipeptide-repeat-protein toxicity** *Nat. Genet* **50**:603–612

22. Stewart M. (2007) **Molecular mechanism of the nuclear protein import cycle** *Nat. Rev. Mol. Cell Biol* **8**:195–208

23. Strambio-De-Castillia C., Niepel M., Rout M.P. (2010) **The nuclear pore complex: bridging nuclear transport and gene regulation** *Nat. Rev. Mol. Cell Biol* **11**:490–501

24. Kalita J., Kapinos L.E., Lim R.Y.H. (2021) **On the asymmetric partitioning of nucleocytoplasmic transport - recent insights and open questions** *J. Cell Sci* **134**

25. Popken P., Ghavami A., Onck P.R., Poolman B., Veenhoff L.M. (2015) **Size-dependent leak of soluble and membrane proteins through the yeast nuclear pore complex** *Mol. Biol. Cell* **26**:1386–1394

26. Timney B.L., et al. (2016) **Simple rules for passive diffusion through the nuclear pore complex** *J. Cell Biol* **215**

27. O'Reilly A.J., Dacks J.B., Field M.C. (2011) **Evolution of the karyopherin-β family of nucleocytoplasmic transport factors; ancient origins and continued specialization** *PLoS One* **6**

28. Görlich D., Seewald M.J., Ribbeck K. (2003) **Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation** *EMBO J* **22**:1088–100

29. Pumroy R.A., Cingolani G. (2015) **Diversification of importin-α isoforms in cellular trafficking and** *Biochemical Journal* **466**:13–28

30. Lott K., Cingolani G. (2011) **The importin β binding domain as a master regulator of nucleocytoplasmic transport** *Biochim Biophys Acta* **1813**:1578–92

31. Nanaura H., et al. (2021) **C9orf72-derived arginine-rich poly-dipeptides impede phase modifiers** *Nat Commun* **12**

32. Jafarinia H., Van der Giessen E., Onck P.R. (2022) **Molecular basis of C9orf72 poly-PR interference with the β-karyopherin family of nuclear transport receptors** *Sci. Rep* **12**

33. Ryan S., Rollinson S., Hobbs E., Pickering-Brown S. (2022) **C9orf72 dipeptides disrupt the nucleocytoplasmic transport machinery and cause TDP-43 mislocalisation to the cytoplasm** *Sci. Rep* **12**

34. Zhang K., et al. (2018) **Stress Granule Assembly Disrupts Nucleocytoplasmic Transport** *Cell* **173**:958–971

35. Jafarinia H., van der Giessen E., Onck P.R. (2020) **Phase Separation of Toxic Dipeptide Repeat Proteins Related to C9orf72 ALS/FTD** *Biophys. J* **119**:843–851

36. Semmelink M.F.W., et al. (2022) **Nuclear transport under stress phenocopies transport defects in models of C9Orf72 ALS** *bioRxiv*

37. Pumroy R.A., Ke S., Hart D.J., Zachariae U., Cingolani G. (2015) **Molecular determinants for nuclear import of influenza A PB2 by importin α isoforms 3 and 7** *Structure* **23**:374–384

38. Cook A., et al. (2005) **The structure of the nuclear export receptor Cse1 in its cytosolic state reveals a closed conformation incompatible with cargo binding** *Mol. Cell* **18**:355–367

39. Renault L., et al. (1998) **The 1.7 Å crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller** *Nature* **392**:97–101

40. Hillig R.C., et al. (1999) **The crystal structure of rna1p: a new fold for a GTPase-activating protein** *Mol. Cell* **3**:781–791

41. Bullock T.L., Clarkson D.W., Kent H.M., Stewart M. (1996) **The 1.6 Å resolution crystal structure of nuclear transport factor 2 (NTF2)** *Journal of molecular biology* **260**:422–431

42. Bennion Callister J., Ryan S., Sim J., Rollinson S., Pickering-Brown S.M. (2016) **Modelling C9orf72 dipeptide repeat proteins of a physiologically relevant size** *Hum. Mol. Genet* **25**:5069–5082

43. Swaminathan A., et al. (2018) **Expression of C9orf72-related dipeptides impairs motor function in a vertebrate model** *Hum. Mol. Genet* **27**:1754–1762

44. White M.R., et al. (2019) **C9orf72 Poly(PR) Dipeptide Repeats Disturb Biomolecular Phase Separation and Disrupt Nucleolar Function** *Mol. Cell* **74**:713–728

45. Higurashi M., Ishida T., Kinoshita K. (2009) **PiSite: a database of protein interaction sites using multiple binding states in the PDB** *Nucleic Acids Res* **37**:D360–4

46. Haberland J.r., Becker J.r., Gerke V. (1997) **The acidic C-terminal domain of rna1p is required for the binding of Ran· GTP and for RanGAP activity** *J. Biol. Chem* **272**:24717–24726

47. Seewald M.J., Körner C., Wittinghofer A., Vetter I.R. (2002) **RanGAP mediates GTP hydrolysis without an arginine finger** *Nature* **415**:662–666

48. Ghavami A., Veenhoff L.M., Van Der Giessen E., Onck P.R. (2014) **Probing the disordered domain of the nuclear pore complex through coarse-grained molecular dynamics simulations** *Biophys. J* **107**:1393–1402

49. Ananth A.N., et al. (2018) **Spatial structure of disordered proteins dictates conductance and selectivity in nuclear pore complex mimics** *eLife* **7**:1–24

50. Fragasso A., et al. (2021) **A designer FG-Nup that reconstitutes the selective transport barrier of the nuclear pore complex** *Nat Commun* **12**

51. Ghavami A., van der Giessen E., Onck P.R. (2013) **Coarse-Grained Potentials for Local Interactions in Unfolded Proteins** *J. Chem. Theory Comput* **9**:432–440

52. Ghavami A., van der Giessen E., Onck P.R. (2016) **Energetics of Transport through the Nuclear Pore Complex** *PLoS One* **11**:e0148876–e0148876

53. Dekker M., van der Giessen E., Onck P.R. (2022) **Liquid-liquid phase separation of intrinsically disordered FG-Nups is driven by highly-dynamic hydrophobic FG-motifs** *bioRxiv*

54. Jumper J., et al. (2021) **Highly accurate protein structure prediction with AlphaFold** *Nature* **596**:583–589

55. Varadi M., et al. (2022) **AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models** *Nucleic acids research* **50**:D439–D444

Article and author information

Hamidreza Jafarinia
Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747AG Groningen, The Netherlands
ORCID iD: 0000-0003-1205-2978

Erik Van der Giessen
Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747AG Groningen, The Netherlands
ORCID iD: 0000-0002-8369-2254

Patrick R. Onck
Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747AG Groningen, The Netherlands
For correspondence: p.r.onck@rug.nl
ORCID iD: 0000-0001-5632-9727

Copyright

Editors

Reviewing Editor
Donald Hamelberg
Georgia State University, Atlanta, United States of America

Senior Editor
Qiang Cui
Boston University, Boston, United States of America

Reviewer #1 (Public Review):

Jafarinia et al. have made an interesting contribution to unravel the molecular mechanisms underlying pathological phenotypes of repeat expansion of the C9orf72 gene.

The repeat expression leads to expression of polyPR proteins. Using coarse-grained molecular dynamics simulations, the authors identify putative binding partners involved in nucleocytoplasmic transport (NCT), and conjecture that polyPR affects essential processes by binding to NCT-related proteins.

The results are well-reported, but only putative, and need experimental support to be more conclusive. Also, comparison with results from all-atom MD simulations in explicit water could help verify the results. But even without these, the work is very useful as a first step to unravel the role of polyPR and related peptides.

- https://doi.org/10.7554/eLife.89694.2.sa1

Reviewer #3 (Public Review):

Summary:
Onck and co-workers present in this work the identification of binding partners and sites of polyPR on various nuclear transport components and elucidate how polyPR might potentially influence the transport process. It's interesting to note that some interaction sites on transport components also serve as their inherent/functional binding sites (Figure 3). The difference in the effects between short polyPR (PR7) and long polyPR (PR50) is also evident, although the authors might need to clarify the mechanisms better. Overall, I find this manuscript well organized and concisely written, and it would greatly enhance our understanding of the toxicity induced by polyPR.

Strengths:
The 1-bead per atom force field model used in the study is well-tuned for studying the interactions between polyPR and proteins, as the essential cation-pi interactions (between Arg and Phe/Tyr/Trp) was included using a 8-6 LJ model.

Weaknesses:
To cite the author's response: "At the moment, accurately capturing the binding of NCT components to their native binding targets and the competition with polyPR are best resolved by all-atom molecular dynamics simulations, which come with significant computational demands. This level of detail and computation-intensive analyses is beyond the scope of the current study."

- https://doi.org/10.7554/eLife.89694.2.sa0

Author Response

The following is the authors' response to the original reviews.

We thank the Editor and the referees for their questions and remarks. In this document we provide a point-by-point response to revisions requested by the reviewers.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public Review):***
>
> *Jafarinia et al. have made an interesting contribution to unravelling the molecular mechanisms underlying pathological phenotypes of repeat expansion of the C9orf72 gene. The repeat expression leads to the expression of polyPR proteins. Using coarse-grained molecular dynamics simulations, the authors identify putative binding partners involved in nucleocytoplasmic transport (NCT), and that conjecture that polyPR affects essential processes by binding to NCT-related proteins. The results are well-reported, but only putative, and need experimental support to be more conclusive. Also, a comparison with results from all-atom MD simulations in explicit water could help verify the results. But even without these, the work is very useful as a first step to unravel the role of polyPR and related peptides.*

We greatly appreciate the reviewer's positive assessment of our work and the suggestions. We acknowledge the need for more experimental validation of the binding behavior of some of the transport components. Our results coincide with the experimental findings of Hutten et al. [1] ([16] in our paper) for example regarding the binding of polyPR to Kapβs and Imp□s, but experimental validation of additional transport components, especially for RanGAP, would be valuable. We hope that our work will inspire colleagues from the field to actually perform such experiments.

We also agree with the reviewer's suggestion that all-atom simulations can provide further details on the molecular conformations at the local NTR-PR binding regions. Nonetheless, such simulations for all transport components, particularly for interactions involving large conformational flexibility of longer polyPR chains such as PR50, would require significant computational expenses. In a recent publication (Jafarinia et al. [2]) we reported on the close resemblance in binding behavior between our coarse-grained MD data and the all-atom MD simulations of (Nanaura et al. [3]), both showing polyPR binding to a negatively-charged cavity of Kapβ2. We expect future MD simulations to elucidate more atomistic detail with the continuously increasing power of high-performance computing clusters.

> ***Reviewer #2 (Public Review):***
>
> *This study used coarse-grained molecular dynamics simulation to explain how the binding of polyPR might interfere with distinct stages of the transport cycle. This finding shows that the interaction between polyPR and transport components is driven by electrostatic interactions and is correlated with the salt concentration and the length of polyPR, providing an important basis for subsequent exploration of the impact of C9orf72 R-DPRs on NCT disruption.*

We appreciate the reviewer's positive feedback and the recognition of the significance of our work.

> ***Reviewer #3 (Public Review):***
>
> *Onck and co-workers present in this work the identification of binding partners and sites of polyPR on various nuclear transport components and elucidate how polyPR might potentially influence the transport process. It's interesting to note that some interaction sites on transport components also serve as their inherent/functional binding sites. The difference in the effects between short polyPR (PR7) and long polyPR (PR50) is also evident, although the authors might need to clarify the mechanisms better. Overall, the manuscript is well organized and concisely written, and it would greatly enhance our understanding of the toxicity induced by polyPR. In general, the 1-bead per atom force*

> *field model used in the study is well-tuned for studying the interactions between polyPR and proteins, as the essential cation-pi interactions (between Arg and Phe/Tyr/Trp) were included using an 8-6 LJ model.*

We thank the reviewer for recognizing the suitability of our 1-bead-per-amino-acid force field for studying R-DPRs' interactions with transport components and for acknowledging our work's contribution to understanding polyPR toxicity mechanisms. Below we comment on the mechanisms describing the difference between short and long polyPR molecules.

> ***Recommendations for the authors:***
>
> 1. *Regarding Figure 2 (also see below for more specific comments), there is a major concern that the dipole moment is not included in Fig 2b (as the correlation is better with f=0), but the authors still conclude that this is generally important (lines 258-261). As a minimum, this needs to be discussed more carefully. Is f (i..e. the importance of dipole moment for binding) dependent on the specific binding partner, or what is going on? Maybe, there is a good explanation?*

Indeed, the significance of the dipole moment depends on the specific type of transport component involved. Our analysis reveals that for Kapβs, see figure 2b, the best-fit is obtained with f=0, indicating that the separation of charge within Kapβs has a relatively minor effect on their interaction with polyPR. Instead, the primary determinant for polyPR-Kapβ interaction appears to be the net charge per residue (NCPR), with a more negative NCPR leading to stronger interactions.

We attribute this behavior to the structural characteristics of Kapβs, particularly the superhelical structure which features inner and outer surfaces with differing charge distributions. Importantly, this structural arrangement creates an inner surface characterized by a negative electrostatic potential. As demonstrated in our previous work, polyPR predominantly binds to this negatively charged cavity within Kapβs. Consequently, the separation of charges on the Kapβ surface becomes less influential compared to the overall charge. Other transport components, however, depicted in figure 2a, do not share this feature and the distribution of charges over the surface becomes a more critical factor in polyPR interactions. We have now added this explanation to page 6, and emphasized in the conclusion section that the effect of dipole moment is only observed for the transport components in figure 2a.

> 1. *Write out nucleoporin, Nup, at first appearance (line 51).*

We have changed it in line 51.

> 1. *Fig 1: a (representative) CG structure of polyPR (PR7,PR20 and PR70) would be very useful.*

We have added a CG representation of PR7 and PR20 to figure 1.

> 1. *Please use chi-square, not R-square, to evaluate the fit, as chi-square takes experimental errors into account.*

We use R-square as a standard measure to assess the quality of the fit in the simulations, as it considers the summation of residuals. This choice aligns with the methodology we have used in our previous publications and therefore prefer to use this measure here as well.

> *1. Please use a dot (not a full stop) for multiplication in line 151 and Figure 2 legend.*

We made the adjustment in line 151, the caption of figure 2, and the y-axis label of figure S2.

> *1. 330: it is very unconventional to plot half the std dev as an error bar. Please plot the std dev (standard error) of the mean.·*

We made the suggested change and now the error bars in figure 2 are standard errors of the mean (SEM) calculated from block averaging with three blocks at equilibrium. We also amended the caption of figure 2 and the Methods section.

> *1. Please write an explicit equation for the linear relation that is plotted in Figure 2. Something like: C_t = a*(NCPR - f*M/Rg)+b ? That would make it easier to read.*

We have now added the linear equation of the fit to a new table S4, and included a reference to it in the caption of figure 2.

> *1. Fig 2: why is the fit to PR7 not reported/shown?*

The fits for PR7 resulted in R2 values of 0.89 (a) and 0.83 (b) for 200M and of 0.7 (a) and 0.59 (b) for 100 mM. Because of the low R2 values for 100 mM, the fits for PR7 are not shown. We have added this explanation to the caption of figure 2.

> *1. Fig 4: isn't the blue shape KapB (and not importin)?*

We changed "importin" to "Kapβ Imp" for consistency.

> *1. In the interest of reproducibility, a recommendation is to make the scripts for setting up, running, and analyzing the simulations freely available, e.g. at GitHub. This will increase reproducibility and transparency.*

At the moment we do not have the scripts available on GitHub. However, codes can be provided by the authors upon reasonable request, as also mentioned in the data availability statement in the paper.

> *1. Can the authors explain the salient advances in this article versus the one published last year?*

In our previous work, we showed that polyPR binds to the Kapβ family of nuclear transport receptors (NTRs), consistent with experimental findings. While this provided valuable insights, it was essential to broaden our investigation as C9orf72 toxicity not only affects the Kapβ family of NTRs but also disrupts other key regulators of NCT. For instance, recent literature (see lines 87-91 in our paper) showed that Ran and its regulators RanGAP and RanGEF are mislocalized in cells expressing R-DPRs, and genetic screening studies have identified several nucleocytoplasmic transport genes as modifiers of R-DPR-mediated toxicity.

In the present study, we therefore delved deeper into the underlying mechanisms of polyPR-modification of NCT. We focused on exploring whether polyPR directly interacts with Impα isomers, CAS/Cse1, RanGEF, RanGAP, Ran, and NTF2. By doing so, we unveiled a network of direct interactions between polyPR and a remarkably wide range of NCT components. This newfound insight is valuable for interpreting existing experimental findings, such as the

mislocalization of RanGAP. We also demonstrate that polyPR binding is influenced not only by factors such as the net charge per residue and the polyPR chain length, as previously observed for Kapβs, but also by the spatial separation of charges, incorporated by an additional dependence on dipole moments in influencing the total number of contacts with polyPR. This sheds new light on how polyPR interacts with numerous targets within the cellular environment, providing a valuable reference for future (experimental) investigations of R-DPR-compromised nuclear transport. These points are explained in the last paragraph of the introduction and paragraphs 2,3 of the conclusion section. Paragraph 2 of the conclusion is also modified for clarification.

> *1. In Figure 2(a), the vertical coordinates of the first graph do not match the others.*

We have now modified figure 2a left panel to match the others.

> *1. When the polyPR length is large enough, it seems that the binding of polyPR to RanGEF and NTF2 is not significantly improved.*

The binding behavior depends on polyPR length, as well as on the net charge per residue and the dipole moment (expressed as NCPR-fM/R_g). We note that the number of contacts in figure 2 is normalized by the polyPR length so that for both NTF2 and RanGEF the total number of contacts increase with length (PR7 to PR20) when binding occurs. Specifically, for RanGEF, especially at lower ion concentrations (100 mM), PR7 and PR20 exhibit a similar number of contacts per unit length of polyPR. This implies that the absolute number of contacts between PR20 and RanGEF is higher than that of PR7. However, as we extend the polyPR length to PR50, there is a reduction in the number of contacts per unit length of polyPR. This phenomenon indicates that the more extended PR50 has regions that make little to no contact with RanGEF, resulting in a smaller number of contacts per unit length for PR50. Lines 188-195 are now modified to put more emphasis on the difference between number of contacts and number of contacts normalized by polyPR length.

> *1. The representation of the mechanism in Figure 4 is not intuitive enough and the color scheme still needs to be improved.*

We have tried to improve clarity by including the names of each transport component next to their schematic representations.

> *1. Figure 3 shows that the longer polyPR exhibits a higher contact probability with individual residues compared to a shorter polyPR, is this result in conflict with Figure 2?*

We re-iterate here that the number of contacts in figure 2 is normalized by the polyPR length, while the results in Fig. 3 are not.

Figure 3 and figure S4 demonstrate that as the length of polyPR increases, the contact probability of individual residues of transport components for interaction with polyPR also increases.

In figure 2, we have normalized the time-averaged number of contacts by the length of polyPR. For example, in the top-right panel of figure 2a, when comparing results for PR7 with PR50 interaction with RanGAP, a higher value for PR7 indicates that PR7 makes more contacts per unit of its length with RanGAP. In terms of absolute number of contacts, however, the PR50 chain makes more contacts with RanGAP, resulting in a higher contact probability. We now added a sentence (see lines 188-189) for clarification.

In summary, when a short polyPR strongly binds to a transport component (evidenced by a relatively large number of contacts), it makes more contacts per unit length than a large poyPR. This occurs because for shorter polyPRs most of the residues come into contact with the target protein. In contrast, for longer polyPRs, only certain parts of the chain are in contact with the transport components, while other regions make fewer or no contacts. This is explained in lines 188-195.

> *1. In S2 and S3, does the data require an error bar?*

NCPR, defined as total charge divided by sequence length of the transport components, is a constant and therefore figure S3 does not require an error bar.

In figure S3 we have added error bars (standard deviation) for the dipole moment calculated from 2.5 us simulations of the isolated transport components.

> *1. What is the physiological significance when the salt concentration is 100 mM?*

We conducted simulations at two different salt concentrations: 200 mM, which aligns with in vitro conditions as reported in Hutten et al. [1], and a lower 100 mM salt concentration. The inclusion of the 100 mM salt concentration enables us to assess the significance of salt concentration, and to confirm the dominance of electrostatic interactions in polyPR binding. We also note that this range of salt concentration is commonly used in in-vitro experiments [1, 4, 5].

> *1. Please introduce abbreviation NLS in the abstract.*

We added the full name of NLS to the abstract.

> *1. Given the high number of Arg residues in its sequence, polyPR should interact with many proteins. It would be beneficial to discuss the frequency of binding/non-binding interactions of polyPR with nuclear transport components in comparison to general proteins.*

We appreciate the reviewer's comment. While such a comparison is indeed interesting, our study primarily focused on elucidating the interactions between polyPR and crucial nuclear transport components, aiming to provide insights into potential defects in nucleocytoplasmic transport. The broader comparison of polyPR interactions with different protein classes in the proteome is indeed an interesting direction for future research, but out of the scope of the current manuscript.

> *1. The authors should provide a convergence check to determine whether the 2.5 μs simulations are sufficient for sampling the interaction modes, particularly with the long PR50.*

We have included a new figure (figure S5) and additional text in the Methods section to verify that extending the simulation duration does not alter the contact probabilities (which are indicators of binding modes) presented in figure 3a, confirming convergence of our computations.

> 1. *In reference to Figure 4, the upper panel merely summarizes the known transport mechanisms, while the lower part (A-H) provides potential novel insights from this study. Unfortunately, these novel insights are not sufficiently detailed. It is recommended to include more details to make these relevant plots clearer by expanding the corresponding discussions (currently, only the last paragraph in the Results section addresses these). If possible, the authors should also carry out some CG simulations of the most relevant processes to further elucidate the interference caused by polyPR.*

We have taken the reviewer's feedback into consideration and made the suggested revisions. Specifically, we have expanded the last paragraph of the discussion to provide more detailed explanations of the insights derived from our computational model. For each mechanism, we begin by presenting the reader with the baseline understanding of normal function of the transport component. Subsequently, we discuss how the findings presented in figures 2 and 3 offer insights into polyPR's potential interference with the function of NCT components. Furthermore, we have made improvements to the schematic representation of mechanisms in figure 4 to enhance clarity.

At the moment, accurately capturing the binding of NCT components to their native binding targets and the competition with polyPR are best resolved by all-atom molecular dynamics simulations, which come with significant computational demands. This level of detail and computation-intensive analyses is beyond the scope of the current study, but we hope that our results will provide the groundwork for future, more detailed investigations.

References

1. Hutten, S., et al., Nuclear Import Receptors Directly Bind to Arginine-Rich Dipeptide Repeat Proteins and Suppress Their Pathological Interactions. Cell Rep., 2020. 33(12): p. 108538.
2. Jafarinia, H., E. Van der Giessen, and P.R. Onck, Molecular basis of C9orf72 poly-PR interference with the β-karyopherin family of nuclear transport receptors. Sci. Rep., 2022. 12(1): p. 21324.
3. Nanaura, H., et al., C9orf72-derived arginine-rich poly-dipeptides impede phase modifiers. Nat Commun, 2021. 12(1): p. 5301.
4. Brady, J.P., et al., Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. Proceedings of the National Academy of Sciences, 2017. 114(39): p. E8194-E8203.
5. Fisher, R.S. and S. Elbaum-Garfinkle, Tunable multiphase dynamics of arginine and lysine liquid condensates. Nat. Commun., 2020. 11(1): p. 4628.