A Two-amino Acid Mutation Encountered in Duchenne Muscular Dystrophy Decreases Stability of the Rod Domain 23 (R23) Spectrin-like Repeat of Dystrophin.

Sébastien Legardinier, Baptiste Legrand, Céline Raguénès-Nicol, Arnaud Bondon, Serge Hardy, Christophe Tascon, Elisabeth Le Rumeur, Jean-François Hubert

▶ **To cite this version:**

Sébastien Legardinier, Baptiste Legrand, Céline Raguénès-Nicol, Arnaud Bondon, Serge Hardy, et al.. A Two-amino Acid Mutation Encountered in Duchenne Muscular Dystrophy Decreases Stability of the Rod Domain 23 (R23) Spectrin-like Repeat of Dystrophin.. Journal of Biological Chemistry, 2009, 284 (13), pp.8822-32. 10.1074/jbc.M805846200 . inserm-00365879

HAL Id: inserm-00365879

https://inserm.hal.science/inserm-00365879

Submitted on 5 Mar 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A two-amino-acid mutation encountered in a Duchenne muscular dystrophy decreases stability of the R23 spectrin-like repeat of dystrophin

Sébastien Legardinier*, Baptiste Legrand*, Céline Raguénès-Nicol, Arnaud Bondon, Serge Hardy[§], Christophe Tascon, Elisabeth Le Rumeur and Jean-François Hubert.
* contributed equally to this work.
UMR CNRS 6026 Interactions cellulaires et moléculaires, équipe RMN et Interactions Lipides-Protéines. [§] UMR CNRS 6061 Génétique et développement.
Université Rennes1. IFR 140. Faculté de Médecine, CS 34317, 35043 Rennes Cedex, France.

Running title: A DMD mutation destabilizing dystrophin spectrin-like repeat.

Corresponding author: Prof. Jean-François Hubert, Université Rennes1, UMR CNRS 6026, RMN-ILP, IFR 140, Faculté de Médecine, CS 34317, 35043 Rennes Cedex, France.
Tel: (33) 2 2323 6115; Fax: (33) 2 2323 4606; E-mail: jfhubert@univ-rennes1.fr

Lack of functional dystrophin causes severe Duchenne muscular dystrophy (DMD). While dystrophin's native function is still unknown, its subsarcolemmal location and association with the cytoskeleton and membrane-spanning entities suggests a role in the mechanical regulation of membrane stress during muscle contraction-relaxation cycles. Phenotype rescue by expression of truncated forms of dystrophin in dystrophin deficient *mdx* mice has shown that some parts of the rod-shaped domain of 24 spectrin-like repeats could be essential. We compared biophysical properties of the repeat 23 (R23) of human dystrophin rod domain and one double mutant in helix C, E2910V-N2912D encountered in a DMD. Basal CD and tryptophan fluorescence spectra were superimposable for both proteins while NMR spectra revealed mild change in tryptophan environment. Thermal denaturation analysis by CD showed that wild-type protein remains folded at 37°C, while the mutant is partly unfolded. Urea unfolding analysed by tryptophan fluorescence showed a decrease of mid-denaturing urea concentration for the mutant. Stopped-flow spectrophotometry showed that urea denatured WT R23 refolded much faster than the mutant. On the basis of 3D structures built by homology molecular modelling, we discuss the modifications of stability induced by the double mutation. Taken together, our results show that a two-amino acid mutation in repeat 23 of dystrophin rod domain, causing DMD, induces a significantly decreased stability of R23 repeat that may participate in mechanically weakened dystrophin-deficient muscle. These results represent a contribution to the functional dissection of the dystrophin protein, thus emphasizing the role of the dystrophin rod domain in muscle cell.

Dystrophin-associated muscular dystrophies range from the severe Duchenne muscular dystrophy (DMD) to the milder Becker muscular dystrophy (BMD). Molecular genetic studies indicate that both disorders are the result of mutations in the huge gene that encodes dystrophin. Analysis of naturally occurring mutations demonstrates that approximately two-thirds of the mutations in DMD and BMD forms are intragenic partial deletions (1). Indeed, deletion of exons disrupting the reading frame leads to premature translational termination and absence of the protein in DMD cases. On the other hand, mutations maintaining the translational reading frame can lead to semi-functional microdystrophins in BMD cases (2,3). The remaining one-third of DMD/BMD patients have rare point mutations, such as micro deletions/insertions or substitutions of one or more nucleotides (4,5). Despite the potential importance of this source of information on protein behaviour in cells, the consequences of these mutations on biological properties of dystrophin are so far poorly documented, mainly due to the size and complexity of the protein's cellular partnership. Dystrophin is a 427 kDa protein included within the multiproteic dystrophin glycoprotein complex (DGC), which spans the sarcolemma and links the actin cytoskeleton with the extracellular matrix of myofibres (6). The N-terminal part of dystrophin is reported as a globular actin binding domain 1 (ABD1) containing two alpha-helical calponin

homology domains linked by a central alpha-helix (7). The C-terminal cysteine-rich domain of dystrophin is the binding region to the DGC through an interaction of its WW domain with a proline-rich motif of the cytoplasmic tail of beta-dystroglycan (8). In this way, dystrophin creates a link between the intracellular cytoskeleton and extracellular matrix, which leads to maintenance of cell integrity of the muscular membrane in regard to contraction and relaxation (6). The dystrophin rod domain is made up of a series of 24 weakly repeating units of 110 amino acids similar to the coiled coil repeats of spectrin, and is disrupted by four potential hinges that may ensure flexibility (9). These homologous sequence motifs have been proposed to form domains with a triple-helical bundle type structure (10,11). Primary sequence analysis and modelling studies have revealed the arrangement of the amino acid sequences of spectrin and dystrophin, which form repeating heptad patterns designated "*a*" to "*g*", as described previously (9,12,13). The hydrophobic "*a*" and "*d*" residues located in the inter-helical region stabilize each triple helical repeating unit adjacent to the helix. On the other hand, "*e*" and "*g*" residues are frequently charged, being located between the inter-helical region and the aqueous phase. Hence, these residues may form inter-helical salt bridges with "*e*" or "*g*" residues of opposite charge in an adjacent helix (12).

To understand fully the biological role of the rod domain, further investigations are needed concerning the relationship between the sequences of the repeating units and their biochemical and biophysical properties. The study of cloned wild type or mutated repeating single units from spectrin-like proteins represents a first significant step forward in the investigation of the rod domain (14-20). Moreover, knowledge has been improved by work on micro and minidystrophins lacking one or several repeats (21-24). The roles of some parts of the potentially flexible rod domain are still largely unknown. Nevertheless, the repeats 11 to 17 of the rod domain have been reported as a second actin-binding domain (ABD2) (23,25). The physiological functional importance of ABD2 remains to be elucidated since normal phenotype is restored in dystrophic mice with minidystrophins lacking the whole ABD2 (22,26). On the other hand, it is noteworthy that the sub-domains R1-3 and R20-24, located between unstructured hinges 1- 2 and 3-4, respectively, are present in the phenotype-restoring minidystrophin Δ17-48 found in a Becker patient (27). However, semifunctional microdystrophin genes have been constructed with as few as four spectrin-like repeats and three hinge regions (22). Additionally, the R1-3 sub-domain has been shown to interact specifically with negative membrane lipids and was defined as lipid binding domain 1 (LBD1) (20). Binding to lipids by this part of the rod domain is strongly mediated through hydrophobic interactions between tryptophan residues in repeat 2 and aliphatic chains of phospholipid fatty acids (16). The proximal R1-3 part of the rod domain could thus be involved in membrane adhesion, contributing to the modulation of membrane deformability in the muscle cell, as suggested by the "curvature sensitivity" of the R2 repeat (28). Lipid binding properties could be associated to shuttling of membrane reservoirs that should prevent rupture of the sarcolemma (29). Conversely, the distal R20-24 region of the rod domain lacks any known functional or binding activity (20). This region is located between hinges 3 and 4 in the rod domain, and it is thought to act as a passive linker or as a structural stabiliser of its adjacent beta dystroglycan binding domain. Its proximity with the beta-dystroglycan could contribute to proper mechanical behaviour of the DGC in muscle cells and ensure membrane conservation during muscle cell contraction and elongation cycles. The actual role of rigidity and flexibility of the whole rod domain, as well as that of individual single repeats, remains largely unknown. The comparative analysis of data from rod domains of spectrin versus dystrophin repeats yields incomplete information due to the fact that spectrin repeats are involved in alpha and beta spectrin heterodimerisation, while no part of the dystrophin rod domain has a known dimer-forming role (30).

Most of the rare BMD point mutations correspond to i) missense mutations in the N-terminal or C-terminal domains, or ii) splice-site mutations that probably act, like BMD deletions, via the production of interstitially deleted in-frame transcripts. Only few missense DMD- or BMD-linked point mutations have been reported in the dystrophin gene and 70 of such mutations were reviewed in the dystrophin gene (31). Among them, the double mutation E2910V-N2912D in exon 59 was associated with a severe DMD (32). In the present study, with the aim of better understanding the role of the rod domain in dystrophin and the molecular basis of

muscular dystrophy, we analysed *in vitro* the consequences of the natural double mutation E2910V-N2912D on isolated single repeat 23 from the rod domain of human dystrophin. We show that mutations induce significant changes in biophysical properties of the repeat, and then compare the homology-based molecular model structures of the wild type and mutated repeat 23 of dystrophin. We propose that these changes in stability of a repeat from the rod domain could explain loss of mechanical properties of dystrophin and thus could explain cell muscle defects.

Experimental procedures

Materials. The pGEX-4T1 plasmid vector, GSTrap^{TM} HP column were purchased from GE Healthcare. The ER2566 bacteria were supplied by Ozyme (St Quentin-en-Yvelines, France) and restriction enzymes by Promega (Charbonières, France).

Cloning. The plasmid pTG11025, kindly provided by Dr S. Braun (Transgene, Strasbourg, France), is an *E. coli* plasmid that carries the cDNA encoding the full-length Dp427m isoform of human dystrophin (NCBI Nucleotide Data base NM-004006). Recombinant dystrophin rod domain wild-type and double mutated single-repeat R23 were cloned downstream of the sequence of Glutathione S-transferase (GST) into pGEX-4T1 vector, using BamH I restriction site of the preScission thrombin recognition sequence and Xho I restriction site, as previously described for wild-type single-repeat R23 (20). Site-directed mutagenesis was carried out by PCR using complementary oligonucleotide primers containing the double missense mutation (E2910V, N2912D) and primers previously used to clone single-repeat R23. The PCR-amplified fragment encoding mutated single-repeat R23 was subcloned into pGEM-T Easy vector (Promega, Charbonnieres, France), sequenced by the dideoxy chain termination method (Genome Express, Meylan, France), and directly cloned in the pGEX-4T1 vector using BamHI and Xho I restriction sites.

Boundaries of wild-type and double mutated single-repeat R23. Boundaries of recombinant R23 single-repeats were extended by assuming the original alignment of spectrin-type triple-helical repeats in dystrophin as proposed previously (13). Both recombinant proteins were extended with 6 additional residues at the N-terminal end including the N-terminal Glycine-Serine dipeptide resulting from thrombin cleavage of GST-fusion proteins and 7 additional residues at the C-terminal end according to Gratzer's group (33,34), who recommended extending the constructs by 6 to 8 residues to improve solubility and stability.

Preparation of wild-type and double mutated single-repeat R23. Plasmid constructs were transformed into the protease-deficient *Escherichia coli* ER2566 strain, and cultures were then grown to an absorbance at 600 nm of 0.5 at 37°C in LB broth supplemented with 100 µg/ml ampicillin. Expression of GST fusion proteins was induced by addition of 1 mM isopropyl-beta-D-thiogalactopyranoside. After 4h, bacteria were harvested by centrifugation at 4000 g for 15 min at 4°C. As previously described (20), pellets from 500 ml culture were resuspended in 20 ml ice-cold lysis buffer (20mM Tris-HCl, pH7.5, 150 mM NaCl) and incubated with 0.5 $mg.ml^{-1}$ lysozyme on ice for 30 minutes before being broken up by sonication (U200S, UKA Labortechnic). After centrifugation at 10 000 g for 20 min at 4°C, filtered clarified extracts were loaded onto a pre-equilibrated 5-ml GSTrap^{TM} HP column at a flow rate of about 1 $ml.min^{-1}$. The column was washed with ten column volumes of lysis buffer at a flow rate of 2 $ml.min^{-1}$, and 250 units thrombin were then added. On-column cleavage took place for 48 hours at 4°C and fractions of 1.5 ml were collected in the washing buffer. Fractions were concentrated on 3K Centricon, and the final purification step consisted of FPLC gel filtration on a Sephacryl S-200 column in lysis buffer containing 0.1 mM EDTA (TNE buffer). Protein purity was assessed by SDS PAGE and concentrations determined by BCA (bicinchoninic acid) protein assay with bovine albumin as standard.

Circular dichroism measurements. Circular dichroism spectra were acquired at 20°C with a path length of 0.2 cm on a Jasco J-715 or 815 spectropolarimeter, equipped with Peltier devices for temperature control. The molar ellipticity was calculated using a 100% α-helix coefficient of -36000 $deg.cm^2.dmol^{-1}$ at 222 nm as described previously (35). The thermal unfolding of each construct was monitored by CD at 222 nm at concentrations of 2.5-5 µM in TNE buffer with temperature increasing 1 deg/min from 15 to

85°C. CD signals were fitted either to two-state or three-state transitions. For the two-state transition, the data are fitted to the following equation:

$$y = \frac{(\alpha_N + \beta_N T) + (\alpha_D + \beta_D T)}{1 + e^{\left(\frac{-\Delta G_{UN}}{RT}\right)}} \qquad \text{(equation 1)}$$

where y is the CD signal at 222 nm at temperature T, α_N and α_D are the intercepts of the native and denatured states, respectively, while β_N and β_D are the slopes of the native and denatured states, respectively, ΔG_{UN} is the free energy of unfolding and R is the gas constant in $cal.mol^{-1}.K^{-1}$. As described previously (36), this equation can be rewritten as equation 2 to facilitate non-linear regression analysis by SigmaPlot 10.0 in Windows:

$$y = (\alpha_N+\beta_N T)/\{1+e^{4Tm(T-Tm)/T\Delta T}\} + (\alpha_D+\beta_D T)/\{1+e^{4Tm(Tm-T)/T\Delta T}\} \qquad \text{(equation 2)}$$

where T_m is the melting temperature and ΔT is the width of the unfolding transition.

In the case of a three-state unfolding process, equation 2 can be written as follows:

$$y = (\alpha_N+\beta_N T)/\{1+e^{4Tm1(T-Tm1)/T\Delta T}\} + (\alpha_I+\beta_I T)/\{1+e^{4Tm1(Tm1-T)/T\Delta T}\} + (\alpha_I+\beta_I T)/\{1+e^{4Tm2(T-Tm2)/T\Delta T}\} + (\alpha_D+\beta_D T)/\{1+e^{4Tm2(Tm2-T)/T\Delta T}\} \qquad \text{(equation 3)}$$

where α_I is the intercept of the intermediate state, β_I the slope of the intermediate state, T_{m1} the melting temperature of the first transition between folded and intermediate state and T_{m2} the melting temperature of the second transition between the intermediate and denatured state.

Steady-state fluorescence measurements and urea unfolding. Tryptophan fluorescence spectra of native or urea-treated proteins at 1µM in TNE buffer were recorded at 295 nm excitation wavelength (bandwidth, 8.4 nm) on a Fluorolog 3 spectrofluorometer (Horiba Jobin-Yvon, France), using 10 x 4-mm quartz cuvettes at 20°C. After appropriate correction, we obtained the emission spectra of non denatured proteins and, for denatured proteins previously incubated for 2 hours in TNE with 0.25 to 8 M urea, we obtained the maximum wavelength (λ_{max}) and intensity of fluorescence from the emission spectra. The unfolded fraction, y, was then plotted as a function of urea concentration, U, and fitted to a one-transition, two-state reversible process as described by (37) as given in the equation:

$$y = \{(\alpha_N+\beta_N U) + (\alpha_D+\beta_D U)\, e^{(mU-mU_{50\%})/RT}\} / \{1+e^{(mU-mU_{50\%})/RT}\} \qquad \text{(equation 4)}$$

where α is the intercept of the native (N) and denatured (D) states, β the slope of the native (N) and denatured (D) states, $U_{50\%}$ is the urea concentration at the mid-point of denaturation, m is the slope of the transition between the native and denatured states, R is the gas constant, 1.987 $cal.mol^{-1}.K^{-1}$ and T the temperature in K. The curves were fitted by non-linear regression with Sigma-plot 10.0 in Windows. The term $mU_{50\%}$ in equation 4 is equal to the free energy of denaturation in the absence of urea, ΔG_{UN}.

In the case of a three-state unfolding process, the equation is as follows:

$$y = (\alpha_N+\beta_N U)/\{1+e^{(mU-mU^a{}_{50\%})/RT}\} - 0.5*(\alpha+\beta_I U)/\{1+e^{(mU-mU^a{}_{50\%})/RT}\} + 0.5*(\alpha_I+\beta_I U)/\{1+e^{(mU^a{}_{50\%}-mU)/RT}\} + 0.5*(\alpha_I+\beta_I U)/\{1+e^{(mU-mU^b{}_{50\%})/RT}\} - 0.5*(\alpha_I+\beta_I U)/\{1+e^{(mU^b{}_{50\%}-mU)/RT}\} + (\alpha_D+\beta_D U)/\{1+e^{(mU^b{}_{50\%}-mU)/RT}\} \qquad \text{(equation 5)}$$

where α_I is the intercept of the intermediate state and β_I the slope of the intermediate state, with $U^a_{50\%}$ and $U^b_{50\%}$ representing the urea concentrations at mid-denaturation for the first and second transitions, respectively.

NMR spectroscopy. All the spectra were recorded at 303 K on a Bruker Avance 500 spectrometer equipped with a triple-resonance cryoprobe (^1H, ^{13}C, ^{15}N). The heteronuclear experiments were conducted on ^{15}N-labelled protein obtained by expression in minimum media in the presence of ^{15}N-labelled ammonium chloride. Purification was performed in the same way as the unlabelled protein. Using standard Bruker sequence, HSQC spectra were recorded in phase-sensitive mode in both dimensions using echo-antiecho TPPI gradient selection. ^1H-^{15}N HSQC spectra were acquired with 16 transients, with spectral windows of 16/32 ppm in the proton/nitrogen dimensions and the carrier set at the water frequency and 118 ppm, respectively. The relaxation delay used was 0.8 s. Spectra were processed with Topspin or NMRpipe (38) and visualized with Topspin or NMRview (39) on a Linux station. A matrix of 2K x 256 points was acquired and zero-filled to 2K x 1K points after apodization by shifted sine-square multiplication and linear prediction in the F1 domain.

Refolding kinetics by stopped-flow spectro-fluorescence. Stopped-flow data were recorded on a BioLogic SFM-3, MOS-250 instrument (Grenoble, France). The dead time of the

stopped-flow device was 2.2 ms. For refolding measurements, stock solutions of 8M urea-treated proteins were prepared for two hours in TNE buffer. The refolding reaction was initiated at 25°C by tenfold dilution of the urea-treated protein in TNE buffer. The final protein concentrations were 0.2-0.5 µM. The time-dependent fluorescence of tryptophan changes were monitored at excitation and emission wavelengths of 295 nm and 345 nm, respectively. To determine rate constants, we used curves derived from the averages of at least ten individual kinetic data points after subtraction of the background urea buffer signal. Results were analysed by Biokine software (BioLogic).

Structure calculation. Individual residues were assigned to the "*a*" to "*g*" positions within the heptads proposed by Winder (13) in predicted hydrophobic repeats. The fold recognition program mGenTHREADER (40,41) was run to detect the best templates for comparative modelling. The alignment proposed by mGenTHREADER is concordant with those proposed by Winder (13) and Kusonoki (36) with respect to the heptad motifs conserved in the spectrin repeats. Secondary structure predictions were performed via PSIPRED (42), incorporating four feed-forward neural networks that analyse the output obtained from PSI-BLAST. Homology models of wild type and mutated R23 proteins were built using MODELLER 9.2 software (43). MODELLER is a program based on satisfaction of spatial restraints generated on the target sequence using its alignment with the 3D structure of the templates. Additional restraints were fixed in compliance with the secondary structure prediction of PSIPRED. The loop modelling protocol of MODELLER 9.2 (44) was used to refine all loops of the selected model. Starting with random conformations, 20 structures were built up for each loop. The generated structures were assessed using MODELLER output as well as additional evaluations: the PROCHECK program was used to assess the stereochemical quality of the structures (45), supplemented by the following profile programs: ProSA-WEB (46), the web-based version of ProSA (47), VERIFY3D (48), ANOLEA (49,50) and by examination on a graphic display. A representative model was selected and its energy minimised with the SYBYL software (TRIPOS Inc., St Louis, MO). The secondary structures of the model were analysed and represented with the Pymol program (Delano Scientific). Hydrophobic and electrostatic potentials on the molecular surfaces were calculated using the rTools (Kristian Rother) and APBS (Adaptive Poisson-Boltzmann Solver) (51) software packages respectively, implemented in Pymol.
http://bioinf.cs.ucl.ac.uk/psipred/
http://www.ncbi.nlm.nih.gov/blast/
http://www.doe-mbi.ucla.edu/Services/Verify_3D
http://swissmodel.expasy.org/anolea/

RESULTS

Protein characterization. Figure 1 shows the amino acid sequence of the R23 repeat of human dystrophin, along with the predicted secondary structure and alignment with the sequence of the repeat 16 from chicken brain α−spectrin R16-17 (PDB:1CUN) (52), which is the template selected for molecular modelling. According to Winder's heptad-based alignment (13), the E>V and N>D mutations in positions 2910 and 2912 of human dystrophin in R23 are located in helix C of the triple helix coiled coil, at positions "*c*" and "*e*" of the respective heptad. Both wild type and mutated R23 were expressed in *E.coli* as described in the experimental procedures, and purified in comparable amounts. N-terminal GST-tags were eliminated on an affinity chromatography column by thrombin cleavage and a second purification step was carried out by size-exclusion chromatography allowing to obtain pure proteins as verified by Coomassie blue stained SDS PAGE gels (Figure 2A). The expected molecular masses of 16764 and 16735 $g.mol^{-1}$ were obtained for wild type and mutated R23 proteins, respectively, by mass spectrometry analysis (INRA Rennes). Analytical size-exclusion chromatography shows that both proteins are monomeric and monodisperse with calculated Stokes' radii of 2.4 nm. We conclude that the double mutation E2910V-N2912D does not induce self aggregation or measurable variation in frictional coefficient of the protein. The secondary structure of both proteins in solution at 20°C was analysed by circular dichroism. The spectra exhibits similar feature (Figure 2B), with canonical helical double minima at 208 and 222 nm. The calculation of secondary structures indicates that wild type and mutated R23 have helical ratios of 83% and 81%, respectively, these values being characteristic of the α-helical

structure of spectrin-like repeats (53). No significant difference as regards shape and maximum emission wavelength can be detected between wild type and mutated R23 in the intrinsic fluorescence spectra of the four tryptophan residues (Figure 2C). Nevertheless, 1D ^{1}H NMR spectra (Figure 3A) show that WT and the mutant exhibit differences in the tryptophan indole proton area around 10 ppm, suggesting a slight change in the chemical environment of one or several tryptophan residues. Modifications around tryptophan residue(s) are also consistent with the changes observed in the strongly upfield-shifted methyl resonances. These shifts are consistent with the close proximity of methyl protons to the aromatic rings of the tryptophan. Overall, the proton spectra for both proteins are characteristic of structured protein species. Amide protons of the protein are spread out over 2 ppm and the several methyl protons are shielded around 0 ppm. The 2D ^{1}H-^{15}N spectra HSQC shown on Figure 3B display a relatively good dispersion for both proteins. On both maps, many cross peaks are found at positions compatible with a large structural similarity between the proteins. However, about 30 % connectivities, are strongly modified when comparing the two spectra.

Unfolding-Refolding. Thermal transition curves of both proteins were followed up by observing changes in CD intensity at 222 nm (Figure 4A). The results suggest that wild-type R23 is denatured in a single-step process with a melting temperature of 66.6°C, while the secondary structure of the R23 E2910V-N2912D mutant is clearly lost in a two-step process with two apparent melting temperatures of 44.5 and 61.6°C, respectively.

The urea denaturation data on Figure 4B show that wild-type protein is denatured in a single-step process with a urea concentration of 4.90 M at mid-denaturation. The mutant exhibits a slight first unfolding phase with a urea mid-denaturation value of 2M, and then a second unfolding transition with a mid-denaturation urea concentration of 4.15 M. Figure 4C shows the refolding kinetics of R23 and R23 E2910V-N2912D monitored by stopped-flow tryptophan fluorescence, these proteins being first unfolded in 8.0 M urea and then rapidly diluted to a final urea concentration of 0.8 M at 25°C. The data were fitted to either one or two exponentials. For both proteins, the experimental curves best fit with single exponentials having rate constants of 56±13 s^{-1} and 26±7 s^{-1} for wild-type and mutated R23, respectively. These values correspond to the means of three independent experiments. Experiments were repeated at 12 and 37°C. In all cases, single exponential data are obtained for both proteins, with significantly lower rate constants for the mutated R23 (data not shown).

Molecular modelling. Structural features of the repeat 23 of human dystrophin. As expected, mGenTHREADER reveals that the best templates are spectrin-repeat containing proteins, particularly repeats 16-17 of the chicken brain α-spectrin (CBα16-17, PDB:1cun)(52) and repeats 8-9 of the human erythroid β-spectrin (HEβ8-9, PDB:1s35) (36). R23 shares 16.2% and 12.7% of sequence identity with CBα16 and HEβ8, respectively. As described previously (36), despite a relative low percentage of sequence identity, structures of individual spectrin-like repeats are nearly identical and provide good models for comparative modelling. The crystal structure of CBα16-17 provided the best 3D-structure template model for R23. Figure 1 presents the sequence alignments of R23 with CBα16. In consistency with PSIPRED predictions and considering the helical continuity in the linker regions between the end of one repeat and the beginning of the adjacent one, we assumed for R23 that the extremities of the A and C helices are fully helical. In consequence, we applied alpha helix restraints on two segments of R23 that are not aligned with the CBα16 sequence, namely the N-terminal region (residues 2799 to 2806) and the N-terminal extremity of the helix C (residues 2893 to 2900), which is predicted as helical by PSIPRED. Forty homology models were built up in this way. The models with GA341 scores higher than 0.6, with the lowest target function and DOPE score were evaluated with ProSA-web and VERIFY3D. In the selected model, the loops were then refined and all final structures were assessed using the strategy described in the experimental procedures. For the most representative result shown in Figure 5, VERIFY3D, ProSA and ANOLEA profiles were indicative of a high quality model with a mean force potential positive for VERIFY3D, a negative one for ANOLEA and ProSA with a ProSA global score of -6.32. After minimization, an excellent quality Ramachandran map is obtained, with no residues in the forbidden zones, 92.3% of residues in the most favoured regions and 6.9% in the allowed zone. The triple

helical coiled-coil fold of R23 is closely similar to those reported for spectrin repeats. The curvature observed in helix B is fully compatible with the presence of the semi-conserved proline residue in position 2860. The principal differences arise from the 20-residues insertion specific to R23, between helices B and C, which forms a long BC loop. The BC loop is longer in R23 than in CBα16, HEβ8 and the other dystrophin repeats, so it does not align with the template. We extensively refined it using the loop module of MODELLER. The best BC loop structures show an alpha helix turn from residue 2880-2884 which is not predicted by PSIPRED.

Inset in Figure 5 shows an enlargement of the area bearing the double mutation. We used MODELLER to build the model of the R23 E2910V-N2912D protein from the R23 model, the conformation of the mutant side chain being optimized by a conjugate gradient and refined using molecular dynamics. The mutated residues are located in positions "*c*" and "*e*" in the third heptad hydrophobic repeat pattern of helix C (see Figure 1). According to the selected model, some hydrophobic residues lie on the inward-facing surface of the helices and likely control stability of R23 in the mutated area. These comprise A2908 and V2911, the "*a*" and "*d*" residues in helix C, which interact in a hydrophobic cluster with the corresponding "*a*" and "*d*" residues in helices A (L2814 and L2917) and B (I2862 and T2865).

The residues mutated in the dystrophic pattern are in heptad positions "*c*" (E2910V) and "*e*" (N2912D), and surround the stabilizing residue "*a*" (V2911) in the heptad. As these two residues are in contact with the aqueous phase, we next used the rTools and the APBS software package to calculate hydrophobic and electrostatic surface scores for the whole R23 WT and mutated repeat. In figure 6A is shown the ribbon plot of the chain fold. An examination of the hydrophobic surface properties, Figure 6B, reveals only a slight change between the WT and the mutated whole repeat in the E2910V substitution area. Electrostatic surface analysis of both proteins is shown on Figure 6C. The blue and red colouring represents the positive and negative mean charge surfaces, respectively. As expected from primary structure analysis, the whole repeat exhibits a global negative charge. In the E2910 area, a slight change in averaged negative charge is observed in the electrostatic surface following the E>V substitution, while a much significant increase of negative surface charge is observed on the opposite face of the repeat in the area corresponding to the location of the N2912D substitution (circled area). To assess the features that might induce repeat stability of the R23 repeat, we then analysed the surface properties of individual helices.

Figure 7 shows single side views of electrostatic surfaces of separated helices of R23; helices are represented in their twisted structure from the model. Uncharged residues, indicated in white, are located on a line along each individual helix (dashed lines). They correspond to the "*a*" and "*d*" residues in heptads forming the hydrophobic interaction core between the three helices. For wild type R23, these models clearly show that each of the three helices exhibits a dipolar structure. Strong complementary electrostatic bonds between opposite mean charges of dipolar antiparallel helices likely contribute strongly to folding and/or stability of the triple helical coiled coil repeat. Interestingly, the electrostatic surface of helix C in the R23 E2910V-N2912D mutant exhibits a significantly increased negative charge in the mutated area when compared to the wild type helix C which exhibit uncharged and hydrophobic residues in this area (dotted frame). This change could explain the loss of stability experimentally determined in the mutated protein.

DISCUSSION

The aim of our work was to understand how point mutations could affect the properties of one repeat of the dystrophin rod domain and lead to a severe Duchenne's muscular dystrophy. We focused on the R23 repeat of the rod domain, which contained a double mutation in a dystrophic patient (31,32) and we evidenced a decreased stability of the repeat of the rod domain bearing the mutation.

Investigations into the biological function of the dystrophin-glycoprotein complex suggest that it has an important mechanical function in stabilizing sarcolemma against stresses imposed during muscle contraction or stretching (54). Dystrophin-deficient muscle exhibits abnormalities in cell function that are not clearly understood and the various aspects of dystrophin function remain poorly known. Nevertheless, its cellular location as well as its association with the actin network and membrane-spanning β-dystroglycan suggest a role in the mechanical coupling of the fibre membrane and intracellular forces (55). The rod domain of dystrophin is

thought to function as a molecular spring or shock absorber, assuming that certain parts of the rod domain function as elastic elements (23,56). The determinants of rigidity and flexibility of the rod domain have to be highlighted. We previously showed that the distal part of the rod domain, R20-24, comprising the repeat 23 exhibited no binding to lipids but very high thermal and urea stability (20).

Wild type and mutant R23, both with flanking peptide extensions at the end of the canonical motifs, were expressed as fusion proteins in *E.coli* and further purified. Following removal of the N-terminal GST tag during the purification step, both proteins were obtained as soluble monomers. The amount of α-helix as measured by circular dichroism is in good agreement with a proper spectrin-like repeat folding. As reported by other authors (33,34), the small sequence extensions likely prevent fraying of the helices at the termini of the isolated repeats, but are unlikely to interact with other residues. At low temperatures, we found nearly identical α-helicities for the wild-type and mutant proteins. This indicates that the valine and aspartic acid side chains present in the mutant (V2910, D2912) are compatible with proper folding of the triple α-helix coiled-coil repeat. We also conclude from hydrodynamic behaviour in size exclusion chromatography that the slight changes observed in hydrophobic and electrostatic surface properties between wild type and mutated whole repeat R23 do not induce *in vitro* self aggregation or proteolysis. We clearly demonstrate that the double E2910V-N2912D mutation induces a strong reduction of thermal and chemical stability. Nevertheless, according to comparative refolding velocity analysis *in vitro*, we can infer that the R23 repeat containing the double mutation might follow a folding process leading to a canonical coiled-coil structure that exhibits a reduced stability.

The amino acid substitutions analysed in the present study do not directly concern the hydrophobic stabilizing coiled-coil residues in positions "*a*" and "*d*" of the characteristic heptads spectrin-like repeats, that are known to induce a dramatic loss of stability of a spectrin single repeat if mutated (53), but those located at the "*c*" (E2910V) and "*e*" (N2912D) positions flanking the stabilizing valine residue in "*d*" heptad position. We can assume that these mutations probably have a limited effect on the R23 global fold. To investigate the structural features of the R23 repeat and the determinants involved in the loss of stability of mutated R23, we have built a homology model and made NMR experiments. Homology modelling allowed us to determine the global structure of the R23 repeat, thus it was possible to identify residues near the mutation site and to speculate about intra or inter-helical interactions that likely control the stability of R23. Since yields of production of highly stable ^{15}N labelled-R23 repeat are important, it was possible to obtain valuable information from NMR experiments. At this stage, we observed that about 30% of the signals have significant chemical shifts variations on the ^{15}N HSQC spectrum, which means that about a third of the mutant R23 residues have their local environment modified compared to the R23 WT. It is interesting to note the impact of the mutation on the spectra obtained, while secondary structures are conserved. To give more insight on the mutation effect, we hope to obtain the assignment of the R23 repeat in order to identify the shifted residues in mutant R23 and if possible determine structures of both proteins. However, assignment of the coiled coil proteins is more difficult than for globular proteins, due to a poorer dispersion of the resonances and the small number of long-range constraints (57). Indeed, preliminary studies (data not shown) on a double-labelled ^{15}N-^{13}C WT R23 show that the Cα and Cβ signals are particularly degenerated and make the resonance of the R23 backbone difficult to assign.

The models obtained by comparative modelling permit us to propose several hypothesis that could explain the difference of stability between the WT and mutated R23. Molecular modelling of the R23 repeat is based on the homology with the chicken brain α-spectrin repeat 16 with respect with the "nested model" (33). As described in the results section, the scores and profiles obtained from the validation softwares are satisfactory. By examination of the model with the Pymol software, we can verify that the global features of a spectrin single repeat are satisfied. The residues "*a*" and "*d*" made hydrophobic interactions, lying on the inward facing surface of the helices. One notable hydrophobic cluster of the R23 coiled coil structure includes the two buried tryptophans W2820, in helix A, and W2915, in helix C. These observations fit well with the previously reported interactions between two conserved tryptophan residues, which represent a stabilizing feature through most of the spectrin-like repeats (58). Among these

canonical conserved tryptophans, W2915 in “a” heptad position appears in the model as the residue closest to the mutated area and more precisely close to the mutated residue D2912 in “e” heptad position while the mutated V2910 in “c” position is located on the opposite face of the coiled coil R23 repeat structure. It is likely that the N2912D mutation predominantly influence the behaviour of W2915, as suggested by NMR analysis. Indeed, the NMR data showed that the chemical environment of one of the four tryptophans was modified in mutant vs. WT. The two other tryptophan residues of R23 (W2807 in helix A and W2925 in helix C) are each located at the extremities of the triple helical coiled-coil structure and thus are unlikely to have an environment that could be significantly affected by the mutations. Consequently, we proposed that the double mutation destabilizes indirectly the hydrophobic cluster comprising these two conserved tryptophans facing each other.

It has been shown that electrostatic interactions participate in structural stability of coiled coil spectrin single repeats (59). We observed that electrostatic potentials at the molecular surfaces of WT and mutant R23 were markedly different at the mutation site of helix C, resulting in differences by considering the whole triple helical repeat. Intra helix electrostatic bonds have been reported to contribute to coiled coil stability (60) In our 3D model, the proximity in helix C of the residue E2910 with K2906 or K2917 may suggest the existence of a stabilizing intrahelical salt bridge, that could be disrupted by the E2910V mutation. More noticeable is our observation that, when analysed individually, the anti-parallel helices A, B and C from the R23 repeat, exhibit opposite electrostatic potentials highlighting the existence of interhelical electrostatic forces between these dipolar charged helices, that could reinforce the stability of the triple helical repeat R23.The apparition of a negative charge in the middle of helix C, while hydrophobic and uncharged in WT R23, could be also involved in the loss of stability of mutant R23.

In summary, the model presented here supports the hypothesis that folding and stability of repeats of dystrophin is driven by hydrophobic as well as long-range electrostatic interactions. Otherwise differences are more important on the electrostatic surfaces than the hydrophobic one between the WT and the mutated R23 models, influenced by the nature of the mutations themselves.

Although the role of the rod domain is still poorly understood, it may ensure the flexibility of dystrophin and thus represent a central function of the dystrophin protein (61). Cooperative effects between spectrin repeats have been described in other proteins belonging to this family (17). Due to structural similarities, such cooperative folding properties might occur in the dystrophin rod domain and thus form part of native molecular events involved in the whole dystrophin molecule. The changes in stability properties following two mutations within R23 as demonstrated here, if transposed to the whole R20-24 domain (20) or to the whole dystrophin protein, might lead to dramatic differences in muscle cell life. In that field, one can consider than the C-terminus part of dystrophin is involved in binding to the membrane glycoprotein complex. These interactions have been characterized at the residue level (62), but may be regulated by other domains within or close to the C-terminus. The R20-24 sub-domain, which includes the R23 repeat studied in the present paper, is located between two hinges and has been shown to exhibit a remarkable high stability but no lipid binding properties (20). The R20-24 subdomain is thus unlikely to act on its own as an elastic shock absorber. Due to the close proximity of this part of the rod domain to the beta-dystroglycan-binding domain of dystrophin, we can hypothesise that beta-dystroglycan binding requires an adjacent domain highly stable by itself or stabilized through interaction with a yet unknown partner.

To conclude, we made use in this study of some of the rare pathogenic missense mutations, with an abundance estimated at less than 2% which are found in the dystrophin gene (31). Our data support a critical role for the rod domain in different aspects of dystrophin function. In addition to linkage to the plasma membrane through lipids and beta-dystroglycan (20), the stability and proper mechanical behaviour of the rod domain appear as fundamental. Decreased stability of R23 repeat may contribute to mechanically weakened dystrophin-deficient muscle. Although the mutations studied here would not affect the integrity of the dystrophin glycoprotein complex, they might modify the mechanical properties of certain parts of the molecule. Analysis of the human wild-type R23 and DMD E2910V-N2912D R23 structures obtained from comparative modelling provides some insights into structural differences between wild-type and DMD forms of dystrophin.

REFERENCES

1. Darras, B. T., Blattner, P., Harper, J. F., Spiro, A. J., Alter, S., and Francke, U. (1988) *Am J Hum Genet* **43**(5), 620-629
2. Koenig, M., Beggs, A. H., Moyer, M., Scherpf, S., Heindrich, K., Bettecken, T., Meng, G., Muller, C. R., Lindlof, M., Kaariainen, H., and et al. (1989) *Am J Hum Genet* **45**(4), 498-506
3. White, S., Kalf, M., Liu, Q., Villerius, M., Engelsma, D., Kriek, M., Vollebregt, E., Bakker, B., van Ommen, G. J., Breuning, M. H., and den Dunnen, J. T. (2002) *Am J Hum Genet* **71**(2), 365-374
4. Chaturvedi, L. S., Mukherjee, M., Srivastava, S., Mittal, R. D., and Mittal, B. (2001) *Exp Mol Med* **33**(4), 251-256
5. Eraslan, S., Kayserili, H., Apak, M. Y., and Kirdar, B. (1999) *Eur J Hum Genet* **7**(7), 765-770
6. Ervasti, J. M., and Campbell, K. P. (1993) *Curr Opin Cell Biol* **5**(1), 82-87
7. Norwood, F. L., Sutherland-Smith, A. J., Keep, N. H., and Kendrick-Jones, J. (2000) *Structure* **8**(5), 481-491
8. Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M. J. (2000) *Nat Struct Biol* **7**(8), 634-638
9. Koenig, M., and Kunkel, L. M. (1990) *J Biol Chem* **265**, 4560-4566.
10. Parry, D. A., Dixon, T. W., and Cohen, C. (1992) *Biophys J* **61**(4), 858-867
11. Speicher, D. W., and Marchesi, V. T. (1984) *Nature* **311**(5982), 177-180
12. Pantazatos, D. P., and MacDonald, R. I. (1997) *J Biol Chem* **272**(34), 21052-21059
13. Winder, S. J., Gibson, T. J., and Kendrick-Jones, J. (1995) *FEBS Lett* **369**(1), 27-33
14. Winograd, E., Hume, D., and Branton, D. (1991) *Proc Natl Acad Sci U S A* **88**(23), 10788-10791
15. Kahana, E., Marsh, P. J., Henry, A. J., Way, M., and Gratzer, W. B. (1994) *J Mol Biol* **235**(4), 1271-1277
16. Le Rumeur, E., Fichou, Y., Pottier, S., Gaboriau, F., Rondeau-Mouro, C., Vincent, M., Gallay, J., and Bondon, A. (2003) *J Biol Chem* **278**, 5993-6001
17. Batey, S., Randles, L. G., Steward, A., and Clarke, J. (2005) *J Mol Biol* **349**(5), 1045-1059
18. An, X., Guo, X., Zhang, X., Baines, A. J., Debnath, G., Moyo, D., Salomao, M., Bhasin, N., Johnson, C., Discher, D., Gratzer, W. B., and Mohandas, N. (2006) *J Biol Chem* **281**(15), 10527-10532
19. Saadat, L., Pittman, L., and Menhart, N. (2006) *Biochim Biophys Acta* **1764**, 943-954
20. Legardinier, S., Hubert, J. F., Le Bihan, O., Tascon, C., Rocher, C., Raguenes-Nicol, C., Bondon, A., Hardy, S., and Le Rumeur, E. (2008) *Biochim Biophys Acta* **1784**(4), 672-682
21. Phelps, S. F., Hauser, M. A., Cole, N. M., Rafael, J. A., Hinkle, R. T., Faulkner, J. A., and Chamberlain, J. S. (1995) *Hum Mol Genet* **4**, 1251-1258
22. Harper, S. Q., Hauser, M. A., DelloRusso, C., Duan, D., Crawford, R. W., Phelps, S. F., Harper, H. A., Robinson, A. S., Engelhardt, J. F., Brooks, S. V., and Chamberlain, J. S. (2002) *Nat Med* **8**, 253-261
23. Rybakova, I., Humston, J., Sonneman, K., and Ervasti, J. (2006) *J Biol Chem* **281**, 9996-10001
24. Wang, B., Li, J., Qiao, C., Chen, C., Hu, P., Zhu, X., Zhou, L., Bogan, J., Kornegay, J., and Xiao, X. (2008) *Gene Ther* **15**(15), 1099-1106
25. Amann, K. J., Guo, A. W., and Ervasti, J. M. (1999) *J Biol Chem* **274**(50), 35375-35380
26. Hanft, L. M., Rybakova, I. N., Patel, J. R., Rafael-Fortney, J. A., and Ervasti, J. M. (2006) *Proc Natl Acad Sci U S A* **103**(14), 5385-5390
27. England, S. B., Nicholson, L. V., Johnson, M. A., Forrest, S. M., Love, D. R., Zubrzycka-Gaarn, E. E., Bulman, D. E., Harris, J. B., and Davies, K. E. (1990) *Nature* **343**(6254), 180-182
28. Le Rumeur, E., Pottier, S., Da Costa, G., Metzinger, L., Mouret, L., Rocher, C., Fourage, M., Rondeau-Mouro, C., and Bondon, A. (2007) *Biochim Biophys Acta* **1768**(3), 648-654
29. Sheetz, M. P., Sable, J. E., and Dobereiner, H. G. (2006) *Annu Rev Biophys Biomol Struct* **35**, 417-434
30. Kahana, E., Flood, G., and Gratzer, W. B. (1997) *Cell Motil Cytoskeleton* **36**(3), 246-252

31. Roberts, R. G., Gardner, R. J., and Bobrow, M. (1994) *Hum Mutat* **4**(1), 1-11
32. Lenk, U., Hanke, R., and Speer, A. (1994) *Neuromuscul Disord* **4**(5-6), 411-418
33. Calvert, R., Kahana, E., and Gratzer, W. B. (1996) *Biophys J* **71**(3), 1605-1610
34. Kahana, E., and Gratzer, W. B. (1995) *Biochemistry* **34**(25), 8110-8114
35. Chen, G. C., and Kane, J. P. (1986) *Methods Enzymol* **128**, 519-527
36. Kusunoki, H., MacDonald, R., and Mondragon, A. (2004) *Structure* **12**, 645-656
37. MacDonald, R., and Pozharski, E. (2001) *Biochemistry* **40**, 3974-3984
38. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J Biomol NMR* **6**(3), 277-293
39. Johnson, B. A., and Blevins, R. A. (1994) *J Biomol NMR* **4**(5), 603-614
40. Jones, D. T. (1999) *J Mol Biol* **287**(4), 797-815
41. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) *Bioinformatics* **16**(4), 404-405
42. Jones, D. T. (1999) *J Mol Biol* **292**(2), 195-202
43. Sali, A., and Blundell, T. L. (1993) *J Mol Biol* **234**(3), 779-815
44. Fiser, A., Do, R. K., and Sali, A. (2000) *Protein Sci* **9**(9), 1753-1773
45. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J Biomol NMR* **8**(4), 477-486
46. Wiederstein, M., and Sippl, M. J. (2007) *Nucleic Acids Res* **35**(Web Server issue), W407-410
47. Sippl, M. J. (1993) *J Comput Aided Mol Des* **7**(4), 473-501
48. Luthy, R., Bowie, J. U., and Eisenberg, D. (1992) *Nature* **356**(6364), 83-85
49. Sippl, M. J. (1990) *J Mol Biol* **213**(4), 859-883
50. Melo, F., and Feytmans, E. (1998) *J Mol Biol* **277**(5), 1141-1152
51. Baker, R. D. (2001) *Lifetime Data Anal* **7**(1), 65-83
52. Grum, V. L., Li, D., MacDonald, R. I., and Mondragon, A. (1999) *Cell* **98**, 523-535
53. Kahana, E., and Gratzer, W. B. (1991) *Cell Motil Cytoskeleton* **20**(3), 242-248
54. Judge, L. M., Haraguchiln, M., and Chamberlain, J. S. (2006) *J Cell Sci* **119**(Pt 8), 1537-1546
55. Claflin, D. R., and Brooks, S. V. (2008) *Am J Physiol Cell Physiol* **294**(2), C651-658
56. McGough, A. (1999) *Curr Biol* **9**(23), R887-889
57. Schnell, J. R., Zhou, G. P., Zweckstetter, M., Rigby, A. C., and Chou, J. J. (2005) *Protein Sci* **14**(9), 2421-2428
58. MacDonald, R. I., Musacchio, A., Holmgren, R. A., and Saraste, M. (1994) *Proc Natl Acad Sci U S A* **91**(4), 1299-1303
59. Djinovic-Carugo, K., Gautel, M., Ylanne, J., and Young, P. (2002) *FEBS Lett* **513**(1), 119-123
60. Kammerer, R. A., Jaravine, V. A., Frank, S., Schulthess, T., Landwehr, R., Lustig, A., Garcia-Echeverria, C., Alexandrescu, A. T., Engel, J., and Steinmetz, M. O. (2001) *J Biol Chem* **276**(17), 13685-13688
61. Bhasin, N., Law, R., Liao, G., Safer, D., Ellmer, J., Discher, B. M., Sweeney, H. L., and Discher, D. E. (2005) *J Mol Biol* **352**(4), 795-806
62. Hnia, K., Zouiten, D., Cantel, S., Chazalette, D., Hugon, G., Fehrentz, J. A., Masmoudi, A., Diment, A., Bramham, J., Mornet, D., and Winder, S. J. (2007) *Biochem J* **401**(3), 667-677

FOOTNOTES

This work is supported in part by the “Association Française contre les Myopathies” (AFM). Sébastien Legardinier and Baptiste Legrand are recipient of grants from the “Conseil Régional de Bretagne”. The authors wish to thank C. Rocher (UMR CNRS 6026), MR Allo and S. Pastezeur (UMR CNRS 6061) for technical assistance and D. Mollé (INRA STLO) for mass spectrometry facilities. M.S.N. Carpenter post-edited the English style.

FIGURE LEGENDS

Fig.1. Sequence alignment of the repeat R16 of chicken brain alpha spectrin (CBα16) and the repeat R23 of human dystrophin (R23 WT). Alignment is represented using ESPript, numbers correspond to the CBα16 sequence. Identical residues are represented as white letters on a red background while similar residues are represented as red letters. The 3 α-helices (HA, HB, HC) for template CBα16 and those predicted by PSIPRED for R23 are represented as squiggles. One heptad is enlarged in the inset, showing the localisation of the E2910V and N2912D mutations in helix C of R23 and their assignment in the heptad pattern.

Fig.2. Comparison of wild type and mutant R23 biochemical properties. ***A.*** SDS-PAGE analysis of purified R23 WT and E2910V-N2912D revealed by Coomassie blue staining. ***B.*** CD spectra of R23 WT (circle) and E2910V-N2912D (triangle) at 2.5 µM. CD spectra were obtained at 20°C in TNE buffer, pH 7.5. ***C.*** Tryptophan fluorescence spectra of R23 WT (circle) and E2910V-N2912D (triangle) at 1 µM in TNE buffer at 20°C. Excitation wavelength is 295 nm with spectra recorded between 310 and 420 nm.

Fig.3. NMR analysis**.** ***A.*** 1D ^{1}H NMR spectra of the R23 WT (bottom) and R23 E2910V-N2912D (top)**.** ***B.*** Overlay of ^{1}H-^{15}N 2D HSQC spectra of the R23 WT (black) and E2910V-N2912D (red).

Fig.4. Unfolding and refolding of wild type and mutant R23. ***A.*** Thermal unfolding as measured by CD spectroscopy at 222 nm, unfolded fraction are plotted as described in methods. R23 WT (circle) follows a single step process (T_m 66.6°C) whereas R23 E2910V-N2912D (triangle) follows a two-step process (T_{m1} 44.5°C and T_{m2} 61.6°C). ***B.*** Urea unfolding as measured by variation of maximum wavelength in tryptophan fluorescence spectra, unfolded fraction are plotted as described in methods. R23 WT (circle) follows a single step process ($U_{50\%}$ 4.9M) whereas R23 E2910V-N2912D (triangle) follows a two-step process ($U^{a}_{50\%}$ 2M and $U^{b}_{50\%}$ 4.2M). ***C.*** Refolding kinetics observed by stopped-flow fluorimetry after a one tenth dilution of R23 WT and R23 E2910V-N2912D, unfolded in 8.0 M urea. Final protein concentrations are of 0.2 µM. Measurements were performed at 25 °C by monitoring fluorescence emission at 345 nm after excitation at 295 nm. Curves show the averages of at least ten individual kinetics. The rate constants k were obtained by fitting exponential functions to stopped-flow data.

Fig.5. Molecular modelling of wild type R23. Pymol representation of the R23 WT global fold. N and C-terminus and helices numbering (HA, HB, HC) are indicated. Important amino-acids are coloured and displayed in stick representation: W2807, W2820, W2915 and W2925 in yellow, E2910 and N2912 in green and P2860 in red. The inset shows, enlarged and rotated, the area of the mutation with charged residues also in colour (blue and red for positive and negative residues respectively).

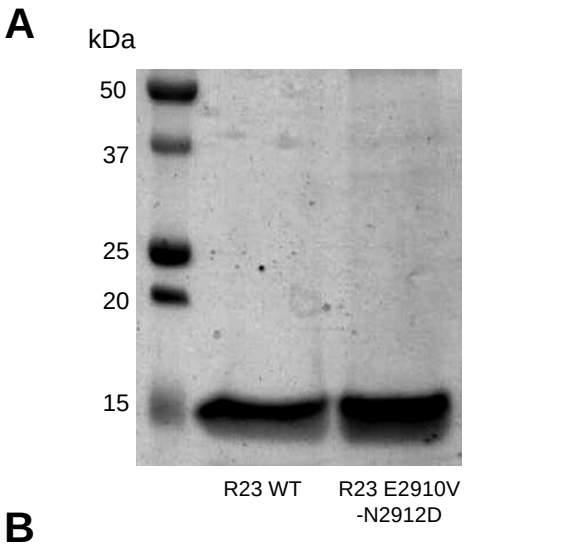
Fig.6. Surface properties of the wild type and mutant R23. ***A***. Orientation of the molecule. ***B.*** Hydrophobic potentials at the molecular surface of R23 WT (left) and R23 E2910V-N2912D (right). Surface is coloured with increasing hydrophobicity from green to red. ***C.*** Electrostatic potentials at the molecular surface of R23 WT (left) and R23 E2910V-N2912D (right). Surface is coloured with increasing electrostatic potentials from red to blue. The areas of main differences between WT and E2910V-N2912D are circled with dotted lines.

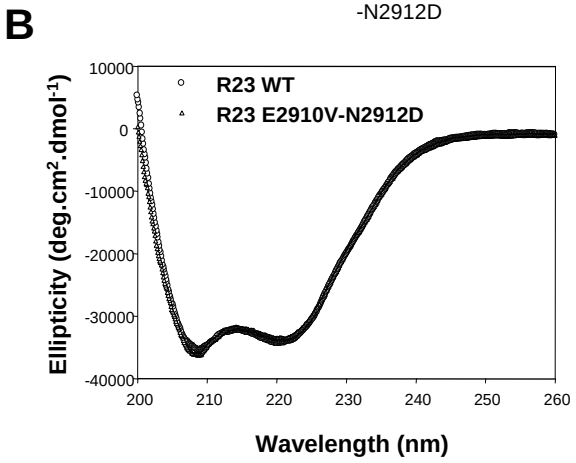
Fig.7. Electrostatic potentials of each individual helix HA, HB and HC for wild type and mutant R23. Surface is coloured with increasing electrostatic potentials from red to blue. Numbers corresponding to first and last amino acids of each helix are mentioned. Dashed lines represent the location of the hydrophobic residues *a* and *d* by which the helices interact. The areas of main differences between WT and E2910V-N2912D are framed with dotted line.

Figure 1

HA
HB
HC
CBα16
R23 WT
2910
2912
AEEVNTE
AEVVDTE
a b c d e f g
R23 WT
R23 E2910V-N2912D

Figure 2


A
kDa
50
37
25
20
15
R23 WT
R23 E2910V
-N2912D


B
10000
0
-10000
-20000
-30000
-40000
200
210
220
230
240
250
260
R23 WT
R23 E2910V-N2912D
Ellipticity (deg.cm2.dmol-1)
Wavelength (nm)

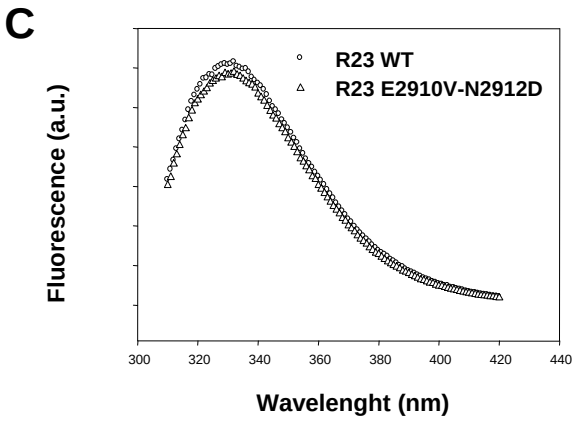

C
300
320
340
360
380
400
420
440
R23 WT
R23 E2910V-N2912D
Fluorescence (a.u.)
Wavelenght (nm)

Figure 3

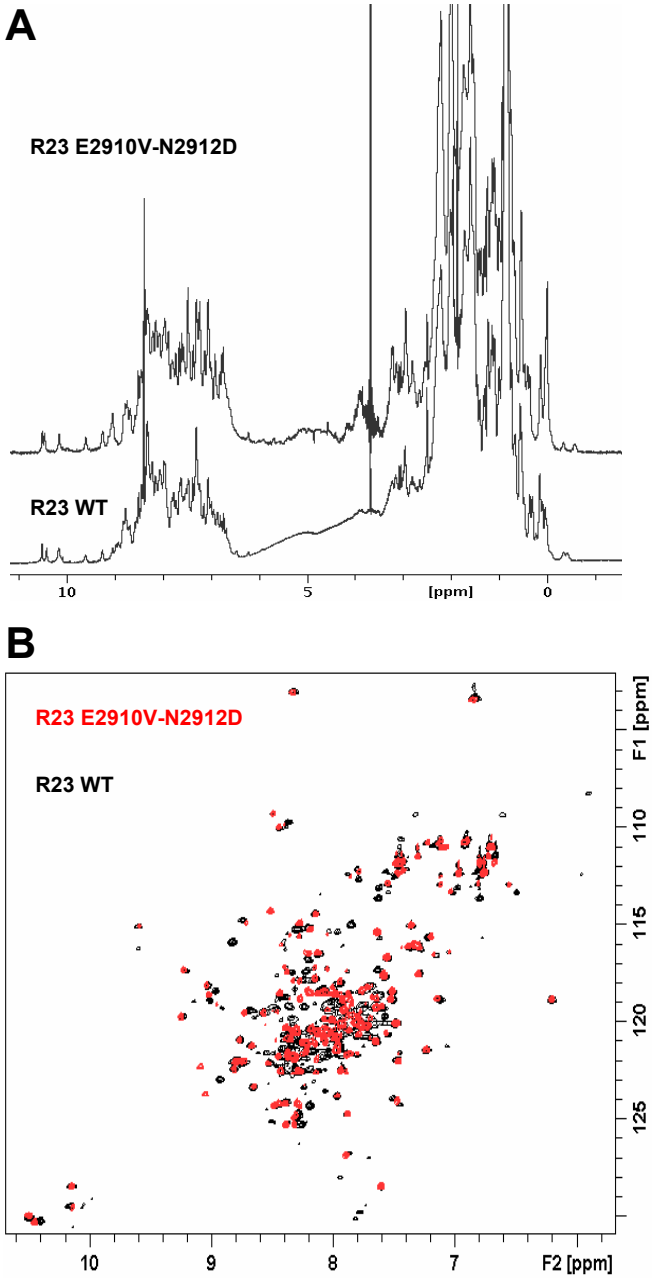

A
R23 E2910V-N2912D
R23 WT
10
5
[ppm]
0
B
R23 E2910V-N2912D
R23 WT
F1 [ppm]
110
115
120
125
10
9
8
7
F2 [ppm]

Figure 4

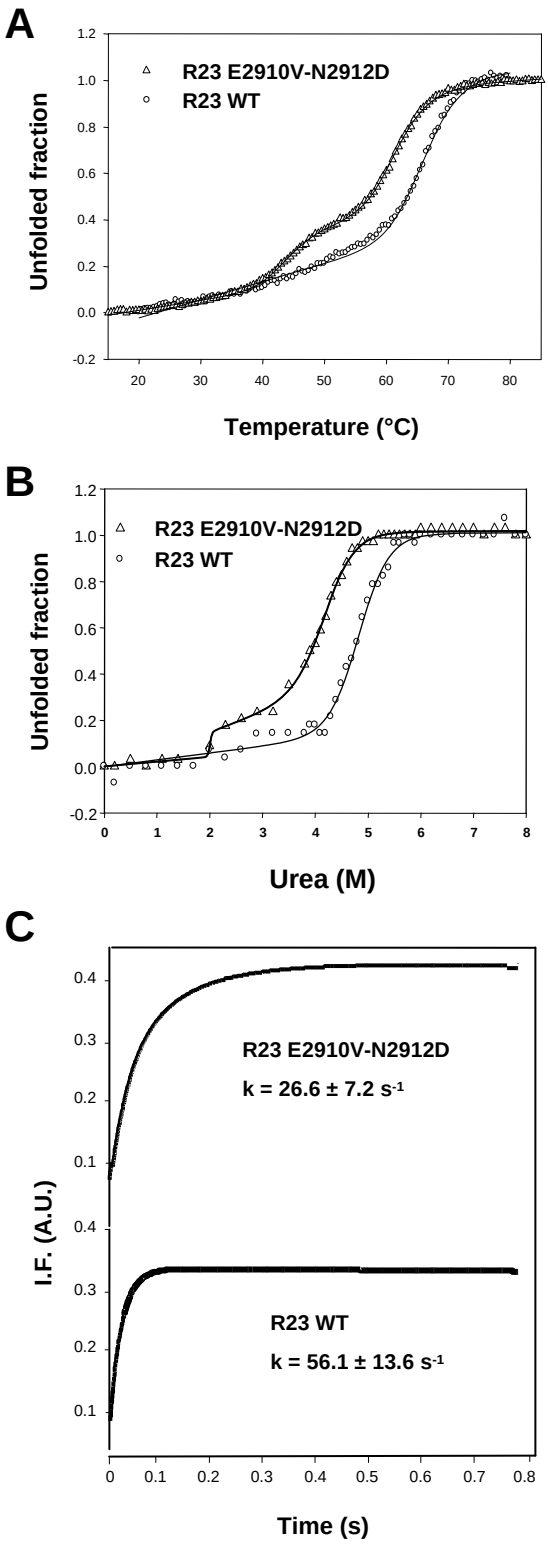

A
R23 E2910V-N2912D
R23 WT
Unfolded fraction
Temperature (°C)
B
R23 E2910V-N2912D
R23 WT
Unfolded fraction
Urea (M)
C
R23 E2910V-N2912D
k = 26.6 ± 7.2 s-1
R23 WT
k = 56.1 ± 13.6 s-1
I.F. (A.U.)
Time (s)

Figure 5

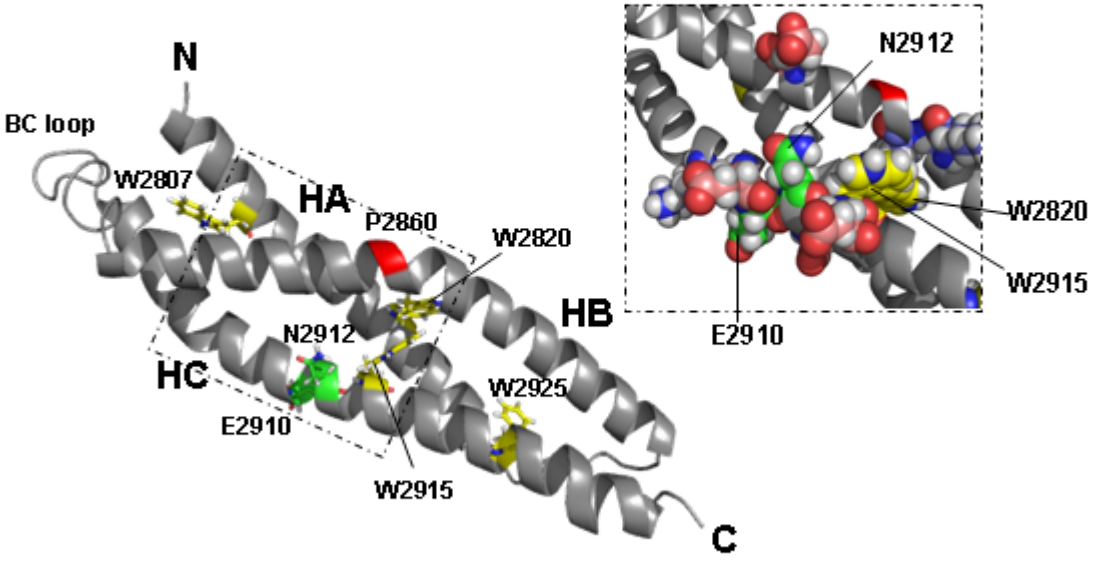
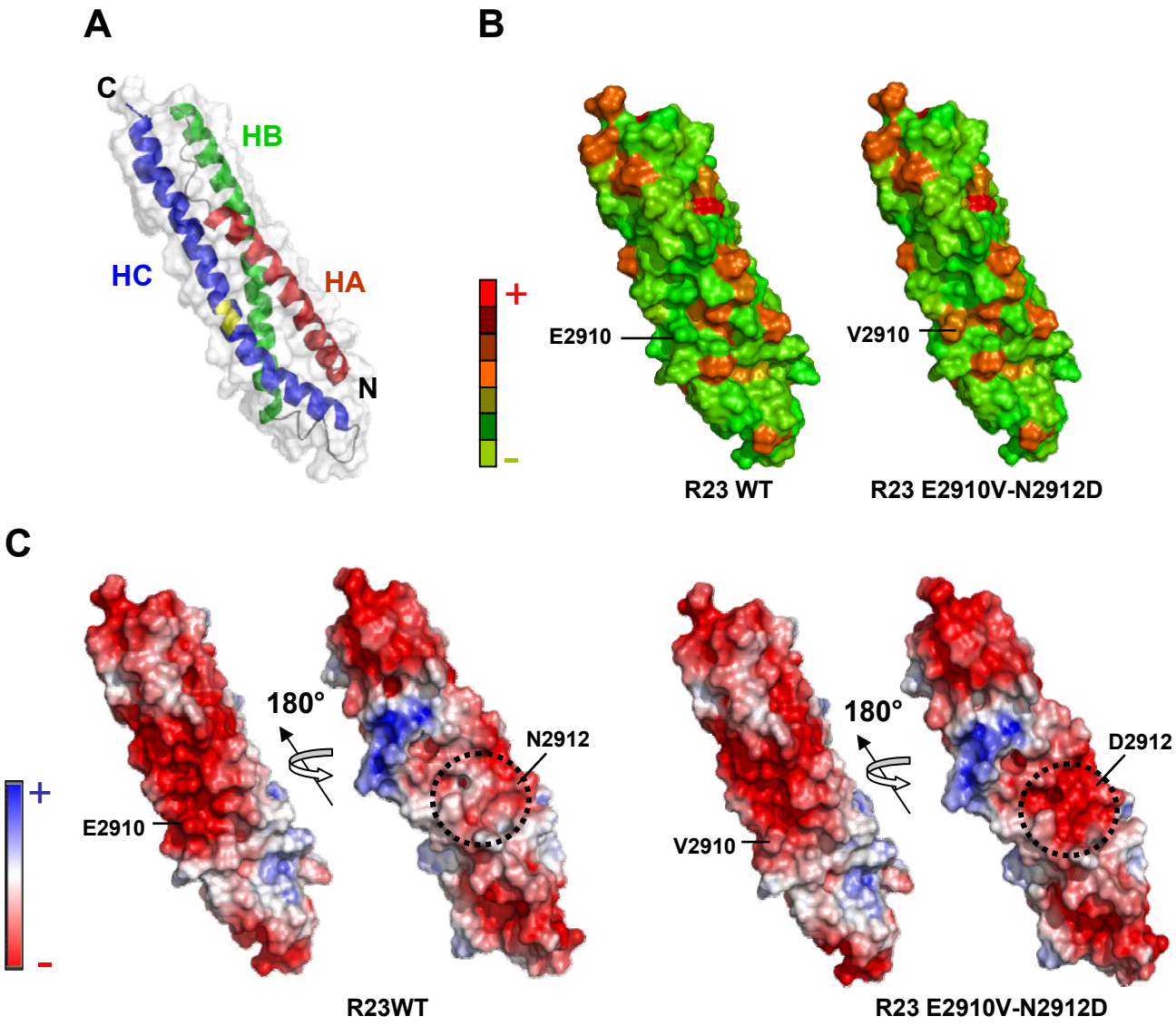

N
BC loop
W2807
HA
P2860
W2820
HB
N2912
HC
E2910
W2925
W2915
C
N2912
W2820
W2915
E2910

Figure 6


A
C
HB
HC
HA
N
B
+
-
E2910
V2910
R23 WT
R23 E2910V-N2912D
C
+
-
180°
E2910
N2912
R23WT
180°
V2910
D2912
R23 E2910V-N2912D

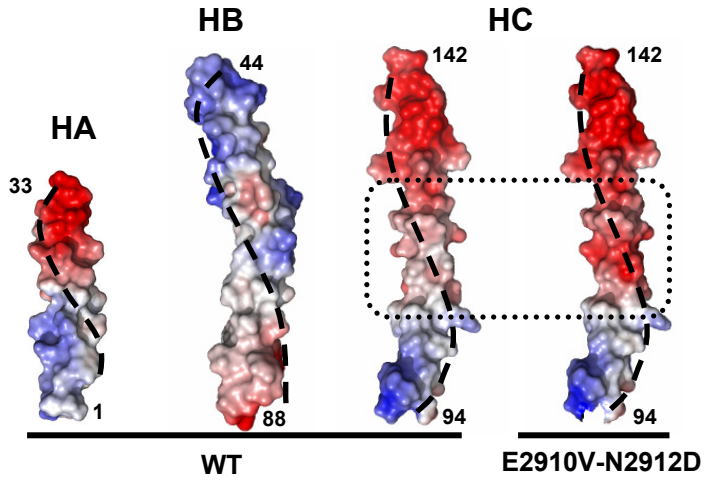



Figure 7