The American Journal of Human Genetics, Volume *104*

Supplemental Data

De Novo Variants Disrupting the HX Repeat Motif of ATN1 Cause a Recognizable Non-Progressive Neurocognitive Syndrome

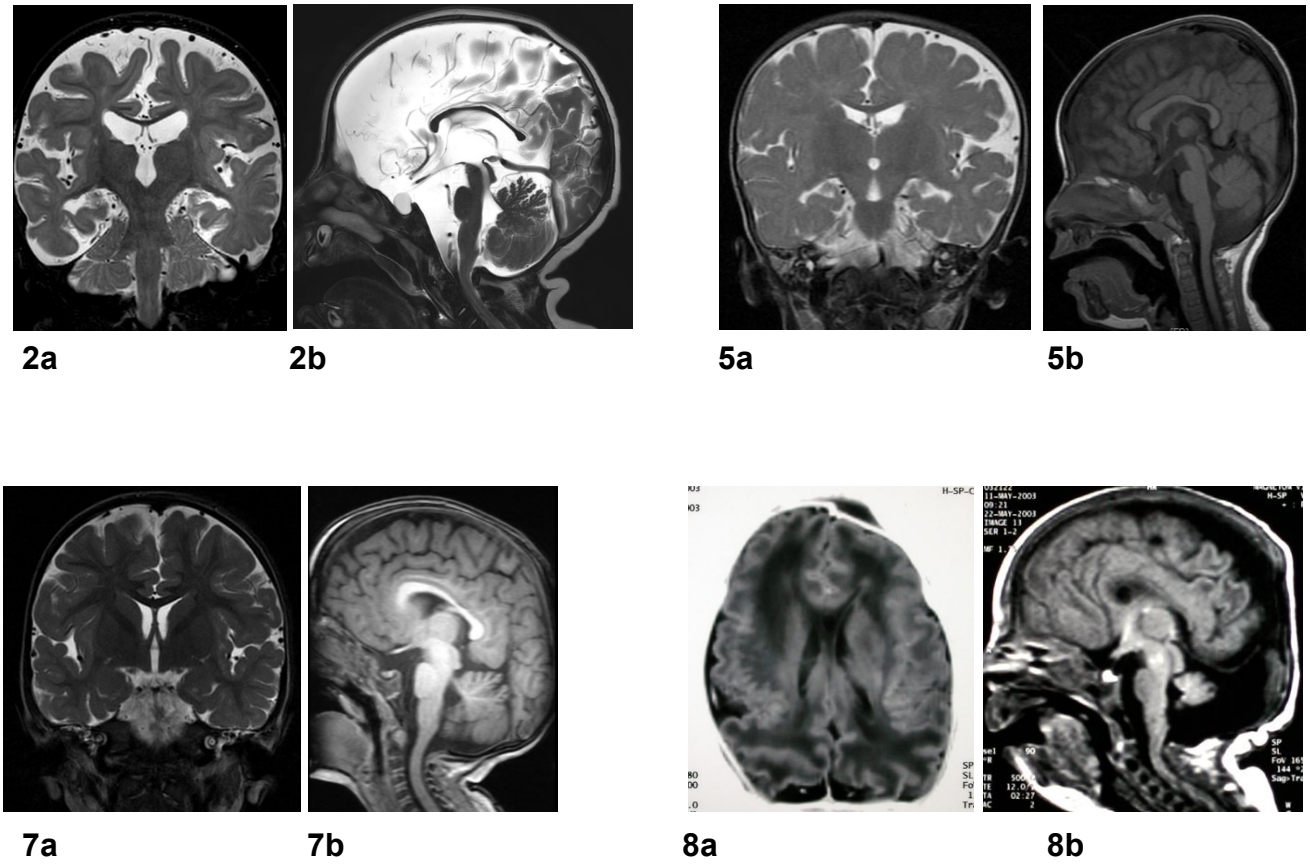
Elizabeth E. Palmer, Seungbeom Hong, Fatema Al Zahrani, Mais O. Hashem, Fajr A. Aleisa, Heba M. Jalal Ahmed, Tejaswi Kandula, Rebecca Macintosh, Andre E. Minoche, Clare Puttick, Velimir Gayevskiy, Alexander P. Drew, Mark J. Cowley, Marcel Dinger, Jill A. Rosenfeld, Rui Xiao, Megan T. Cho, Suliat F. Yakubu, Lindsay B. Henderson, Maria J. Guillen Sacoto, Amber Begtrup, Muddathir Hamad, Marwan Shinawi, Marisa V. Andrews, Marilyn C. Jones, Kristin Lindstrom, Ruth E. Bristol, Saima Kayani, Molly Snyder, María Mercedes Villanueva, Angeles Schteinschnaider, Laurence Faivre, Christel Thauvin, Antonio Vitobello, Tony Roscioli, Edwin P. Kirk, Ann Bye, Jasmeen Merzaban, Łukas Jaremko, Mariusz Jaremko, Rani K. Sachdev, Fowzan S. Alkuraya, and Stefan T. Arold

Supplemental Case Reports

Detailed phenotypic information on all affected individuals is provided in Table S1**.**

Supplemental Figures and Legends

Figure S1:



Legend: Neuroimaging of affected individuals with *de novo ATN1* variants.

Common features between affected individual include thinning of the corpus callosum with low hanging medial parietal lobule which results in anomalous course of the Vein of Galen, polymicrogyria of the sylvian fissure and steep clival angle. Affected individuals exhibit varying degrees of parenchymal atrophy, which is more pronounced in the Sylvian fissures. Of the four individuals where the neuroimaging was available two had clear absence of the falx cerebri and the other two had partial absence.

2a: Coronal T2 MRI of individual 2 illustrating absence of falx cerebri and peri-sylvian polymicrogyria. This patient also has significant parenchymal atrophy with widening of sulci; 2b: Sagittal T2 MRI of individual 2 showing steep clival angle, thin corpus callosum, abnormal course of Vein of Galen and parenchymal atrophy.

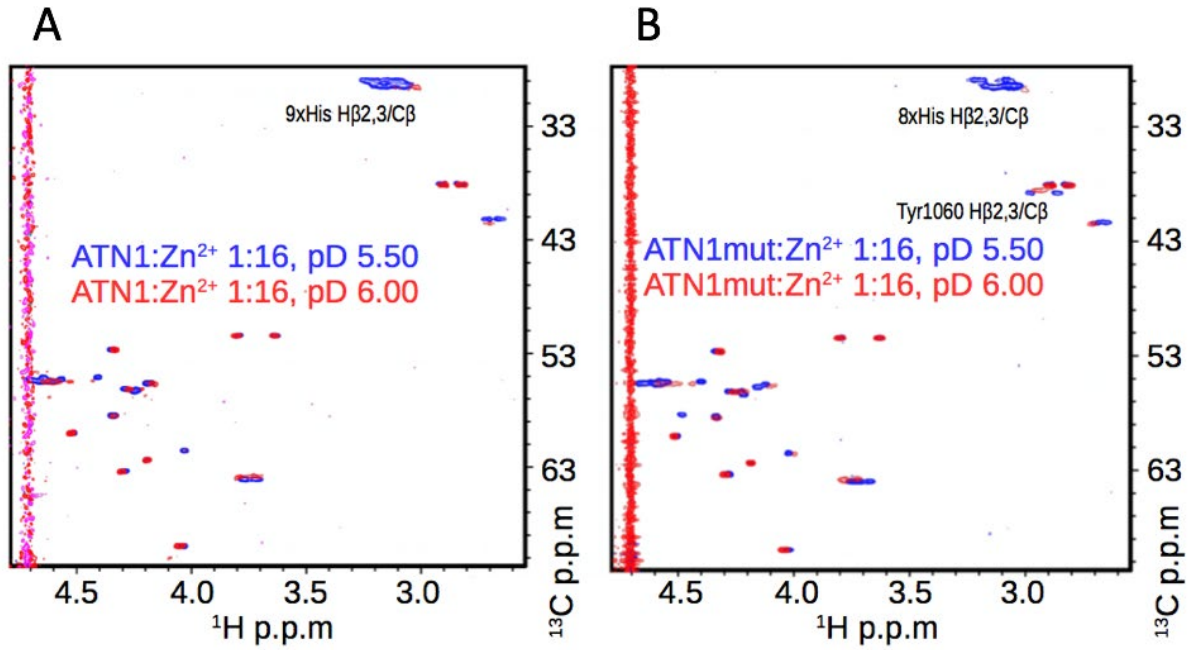
5a: Coronal T2 MRI of individual 5 showing peri-sylvian polymicrogyria, parenchymal atrophy as illustrated by widened Sylvian fissure and sulci and partial absence of falx cerebri;

5b: Sagittal T1 MRI illustrating thinned corpus callosum and abnormal course of Vein of Galen in Individual 5.

7a: Coronal T2 MRI illustrating absence of falx cerebri and peri-sylvian polymicrogyria in individual 7; 7b: Sagittal T1 MRI of individual 7 showing steep clival angle, thin corpus callosum, and abnormal course of Vein of Galen.

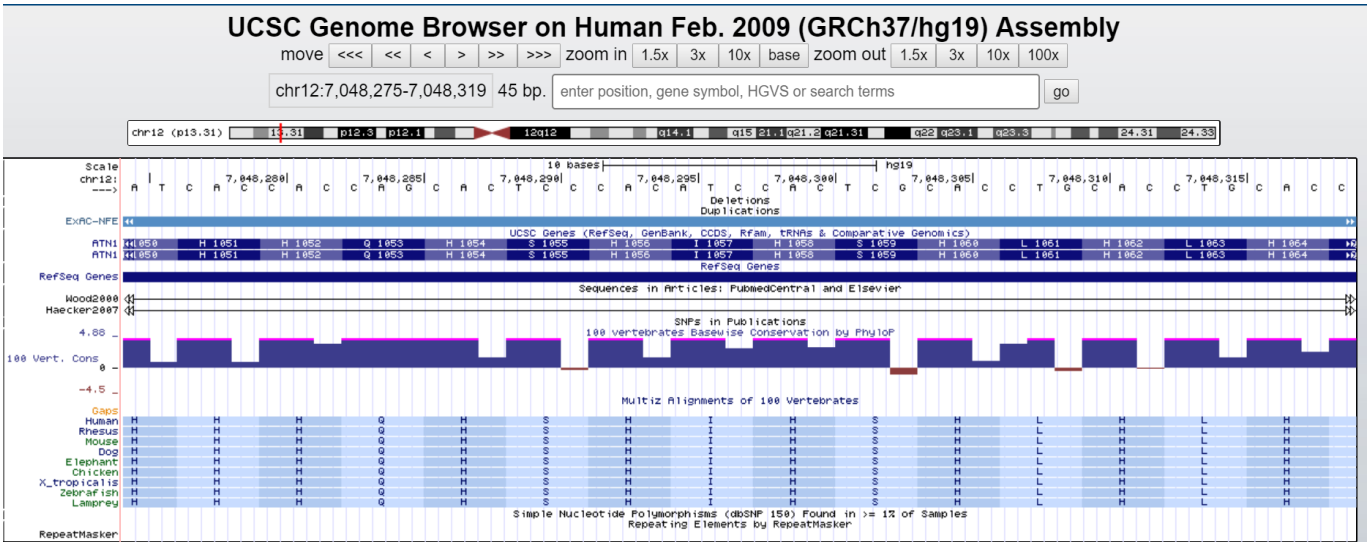
8a: Axial T1 MRI of individual 8 demonstrates polymicrogyria of the right Sylvian fissure; 8b Sagittal T1 MRI of individual 8 demonstrates thin corpus callosum and low hanging medial parietal gyrus. It is difficult to identify the vein of Galen. There is a steep clival angle. This individual exhibits vermian atrophy of the cerebellum.

Figure S2:



Legend: 2D heteronuclear ^{1}H - ^{13}C HSQC correlation spectra. Peptides were recorded using natural abundance of ^{1}H and ^{13}C in synthesized peptides. ATN1: 1046-NVTPHHHQHSHIHSHLHLHQQD-1067 ($ATN1_{1046-1067}$); ATN1mut: 1046-NVTPHHHQHSHIHSYLHLHQQD-1067 ($ATN1_{1046-1067}^{His1060Tyr}$). pD corresponds to the pH in D_2O. A-B: Comparison of Zn^{2+} interaction at pD of 5.5 and 6.0, at a peptide: Zn^{2+} ion ratio of 1:16

Figure S3:



Legend: Evolutionary conservation of the HX domain. Amino acids 150-164 in ATN1 are highly evolutionary conserved across species. Screenshot UCSC Genome Browser.

Supplemental Methods

Next generation sequencing and variant analysis

Individuals 1, 5 and 8 were screened by ES using previously described methodology[1-3].

Sequences of individuals 2, 3, 4, and 6 were captured using the SureSelect Human All Exon V4 (50 Mb), the Clinical Research Exome kit (Agilent Technologies, Santa Clara, CA) or the IDT xGen Exome Research Panel v1.0. Massively parallel (NextGen) sequencing was done on an Illumina system with 100bp or greater paired-end reads. Additional sequencing technology and variant interpretation protocols have been previously described[4].

Individual 7 was screened for sequence and structural variants by WGS using previously described methodology[5; 6] as well for variants in the mitochondrial genome using the in-house pipeline mity: mity has been designed to detect low heteroplasty SNPs and INDELs in the mitochondrial genome from a blood sample.

In all cases, the variants were confirmed and segregation analysis performed by Sanger sequencing using standard methodology**.**

Pattern Search for $(HX)_8$ motifs in the human proteome

Data were obtained using the H-x-H-x-H-x-H-x-H-x-H-x-H-x-H pattern in the PatternSearch program[7]. The initial list of 71 proteins was hand-curated to retain only HX sequences reminiscent to ATN1. Only the HX repeat of each protein sequence is shown.

```
>NP_001007027.1 atrophin-1
----PHHHQHSHIHSHLHLHQ------
>NP_001036146.1 arginine-glutamic acid dipeptide repeats protein isoform a
----PHHHQHSHIHSHLHLHQ------
>NP_001036147.1 arginine-glutamic acid dipeptide repeats protein isoform b
----PHHHQHSHIHSHLHLHQ------
>NP_001098549.2 probable fibrosin-1 *
----FHQHNHQHQHTHQHTHQHF----
>NP_001120703.1 autism susceptibility gene 2 protein isoform 2 *
----FHQHQHQHQHTHQHTHQHT----
>NP_001120729.1 zinc transporter ZIP10 precursor  *
----GHDHSHQHAHGHGHSHGHE------
>NP_001136113.1 fibrosin-1-like protein
----FHQHQHTHQHTHQHTHQHQHT--
>NP_001138356.1 zinc transporter 7
----AHSHDHAHGHGHFHSH-------
>NP_001931.2 atrophin-1
----PHHHQHSHIHSHLHLHQ------
>NP_036234.3 arginine-glutamic acid dipeptide repeats protein isoform a
----PHHHQHSHIHSHLHLHQ------
>NP_056385.1 autism susceptibility gene 2 protein isoform 1
----FHQHQHQHQHTHQHTHQHT----
>NP_065075.1 zinc transporter ZIP10 precursor
----GHDHSHQHAHGHGHSHGHE------
>NP_598003.2 zinc transporter 7
----AHSHDHAHGHGHFHSHD------
>XP_005246746.2 PREDICTED: zinc transporter ZIP10 isoform X1
----GHDHSHQHAHGHGHSHGHE------
>XP_005250314.1 PREDICTED: autism susceptibility gene 2 protein isoform X8
----FHQHQHQHQHTHQHTHQHT----
>XP_005263521.1 PREDICTED: arginine-glutamic acid dipeptide repeats protein isoform X1
----PHHHQHSHIHSHLHLHQ------
>XP_005263523.1 PREDICTED: arginine-glutamic acid dipeptide repeats protein isoform X3
----PHHHQHSHIHSHLHLHQ------
>XP_005266228.1 PREDICTED: fibrosin-1-like protein isoform X3
```

----FHQHQHTHQHTHQHTHQHQHT--
>XP_005266230.1 PREDICTED: fibrosin-1-like protein isoform X4
----FHQHQHTHQHTHQHTHQHQHT--
>XP_005266232.1 PREDICTED: fibrosin-1-like protein isoform X9
----FHQHQHTHQHTHQHTHQHQHT--
>XP_005266234.1 PREDICTED: fibrosin-1-like protein isoform X11
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011509806.1 PREDICTED: zinc transporter ZIP10 isoform X1
----GHDHSHQHAHGHGHSHGHE------
>XP_011509807.1 PREDICTED: zinc transporter ZIP10 isoform X1
----GHDHSHQHAHGHGHSHGHE------
>XP_011509808.1 PREDICTED: zinc transporter ZIP10 isoform X1
----GHDHSHQHAHGHGHSHGHE------
>XP_011509809.1 PREDICTED: zinc transporter ZIP10 isoform X2
----GHDHSHQHAHGHGHSHGHE------
>XP_011514312.1 PREDICTED: autism susceptibility gene 2 protein isoform X1
----FHQHQHQHQHTHQHTHQHT----
>XP_011514313.1 PREDICTED: autism susceptibility gene 2 protein isoform X2
----FHQHQHQHQHTHQHTHQHT----
>XP_011514314.1 PREDICTED: autism susceptibility gene 2 protein isoform X3
----FHQHQHQHQHTHQHTHQHT----
>XP_011514315.1 PREDICTED: autism susceptibility gene 2 protein isoform X4
----FHQHQHQHQHTHQHTHQHT----
>XP_011514316.1 PREDICTED: autism susceptibility gene 2 protein isoform X5
----FHQHQHQHQHTHQHTHQHT----
>XP_011514319.1 PREDICTED: autism susceptibility gene 2 protein isoform X6
----FHQHQHQHQHTHQHTHQHT----
>XP_011514320.1 PREDICTED: autism susceptibility gene 2 protein isoform X7
----FHQHQHQHQHTHQHTHQHT----
>XP_011533105.1 PREDICTED: fibrosin-1-like protein isoform X1
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533106.1 PREDICTED: fibrosin-1-like protein isoform X2
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533107.1 PREDICTED: fibrosin-1-like protein isoform X5
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533108.1 PREDICTED: fibrosin-1-like protein isoform X6
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533110.1 PREDICTED: fibrosin-1-like protein isoform X8
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533112.1 PREDICTED: fibrosin-1-like protein isoform X13
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533114.1 PREDICTED: fibrosin-1-like protein isoform X14
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533118.1 PREDICTED: fibrosin-1-like protein isoform X15
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011539081.1 PREDICTED: zinc transporter 7 isoform X2
----AHSHDHAHGHGHFHSHD------
>XP_011539812.1 PREDICTED: arginine-glutamic acid dipeptide repeats protein isoform X2 [Homo
sapiens
----PHHHQHSHIHSHLHLHQ------
>XP_011544218.1 PREDICTED: probable fibrosin-1 isoform X1
----FHQHNHQHQHTHQHTHQHF----
>XP_011544219.1 PREDICTED: probable fibrosin-1 isoform X2
----FHQHNHQHQHTHQHTHQHF----
>XP_011544221.1 PREDICTED: probable fibrosin-1 isoform X3
----FHQHNHQHQHTHQHTHQHF----
>XP_016855889.1 PREDICTED: zinc transporter 7 isoform X1
----AHSHDHAHGHGHFHSHD------
>XP_016855890.1 PREDICTED: zinc transporter 7 isoform X3
----AHSHDHAHGHGHFHSHD------
>XP_016856847.1 PREDICTED: arginine-glutamic acid dipeptide repeats protein isoform X1
----PHHHQHSHIHSHLHLHQ------
>XP_016856848.1 PREDICTED: arginine-glutamic acid dipeptide repeats protein isoform X1
----PHHHQHSHIHSHLHLHQ------
>XP_016860011.1 PREDICTED: zinc transporter ZIP10 isoform X2
----GHDHSHQHAHGHGHSHGHE------
>XP_016867440.1 PREDICTED: autism susceptibility gene 2 protein isoform X9
----FHQHQHQHQHTHQHTHQHT----

NMR Analysis, extended description

Free ATN1 and ATN1mut in solution

To probe the impact of the patient variants on the molecular behavior of the ATN1 protein we studied two polypeptides containing the $(HX)_n$ motif (n=8) in solution by NMR techniques - one of the native sequence 1046-NVTPHHHQHSHIHS**H**LHLHQQD-1067 (ATN1) and one bearing the pathological variant - with the 1060th histidine exchanged to tyrosine - 1046-NVTPHHHQHSHIHS**Y**LHLHQQD-1067 (ATN1mut). The peptides were synthesized using the solid-phase peptide synthesis (SPPS). Peptides were dissolved in 500 µL of 100% D_2O up to the concentration of 2.2 mg/mL. The pH electrode read-out was manually adjusted to 5.10 (what corresponds to pD 5.51, according to Covington *et al*.,[8]) with the stock solution of 0.1 M NaOD in 100% D_2O. All NMR experiments were performed on the 700 MHz spectrometer at 25ºC. The NMR data was processed by NMRPipe[9] and analyzed with Sparky software (https://www.cgl.ucsf.edu/home/sparky/). Analysis of the ^1H and ^{13}C resonances for the two peptides was done by application of a standard procedure[10] based on the inspection of the 2D homonuclear 2D ^1H-^1H TOCSY (with mixing times 10 and 80 ms) and 2D ^1H-^1H ROESY as well as 2D ^1H-^1H NOESY (with mixing times 300 and 500 ms) supported by the natural abundance 2D ^1H-^{13}C HSQC (separately tuned for the aliphatic and aromatic regions) experiments.

Despite the fact that the primary sequences of peptides contain several HX repetitions the 2D heteronuclear ^1H-^{13}C HSQC correlation spectra of the aliphatic regions show decent signal dispersion **(Fig. 2E)**. This made the residue specific resonance assignment possible either for ATN1 as well as ATN1mut peptide. The peptide main chain $^1H\alpha$, $^{13}C\alpha$ chemical shifts showed that the free peptides adopt highly similar disordered conformation in solution. On the 2D ^1H-^{13}C HSQC spectra of the aromatic region coming from the wild-type ATN1 the cross-peaks corresponding to the imidazole Hδ2/Cδ2 and Hε1/Cε1 atoms of all histidines are clustered around one position - indicating similar averaged local environment as well as dynamical properties for each particular histidine side-chain moiety **(Fig. 2F)**. In contrast to the wild-type ATN1 the Hδ2/Cδ2 and Hε1/Cε1 atoms from the imidazole moieties of histidines within the mutated peptide ATN1mut contribute to the larger dispersion of the cross-peaks of the side-chain aromatic region **(Fig. 2F)**. This indicates that the side-chains from these residues are not "synchronized" thus experience their different local environments and mobility than in case of the wild-type ATN1. To further understand the origins and nature of the observed histidine side-chains "synchronization" phenomenon the 2D homonuclear ^1H-^1H ROESY and 2D ^1H-^1H NOESY (with mixing times 100, 300 and 500 ms) spectra, that provide information

about close contacts (up to 5.50 Å) between the protons through space, were recorded. In line with the low signal dispersion in the aromatic region for the wild-type ANT1 the trivial (|j-i| = 0) and short range (|j-i| < 2) NOE cross-peaks are not present or of substantially lower intensity than those observed for the analogical residues within ATN1mut. This observation indicates an increased side-chain mobility for the residues of ATN1 wild-type in comparison to the ATN1mut. Both peptides precipitate from the solution once pD exceeds the value of 6.00 due to deprotonation of the imidazole rings of histidines.

Titrations of ATN1 and ATN1mut with Zn^{2+} ions

The side-chain imidazole rings of histidines are known to coordinate metal ions [11; 12]. To decipher the impact of the pathological variant on the coordination of the metal ions we monitored the Zn^{2+} ions binding to the ATN1 and ANT1mut peptides by NMR. Measurements with Zn^{2+} ions were performed by the stepwise addition of $ZnCl_2$ stock solution (500 mM in 100% D_2O) to the solution of peptides reaching the peptide:Zn^{2+} molar ratios of 1:0.5, 1:1, 1:4, 1:8, 1:16, 1:32 and 1:48 respectively. For each peptide:Zn^{2+} ratio the pD was checked and corrected with the 0.1 M stock solution of NaOD (if needed) and the same set of 1D and 2D NMR spectra as for the free peptides was recorded.

At pD of 5.50 both peptides ATN1 and ATN1mut bind Zn^{2+} ions via their N-termini as indicated by the chemical shift changes of the N-terminal asparagine $^{1}H\alpha/^{13}C\alpha$ **(Fig. 2G,H)** and $^{1}H\beta2,3/^{13}C\beta$ signals. In line with the "synchronized" behavior of the histidine imidazoles the wild type ATN1 histidines did not bind Zn^{2+} but the ones from ATN1mut do bind as demonstrated by histidine $^{1}H\beta2,3/^{13}C\beta$ chemical shift and signal intensity changes upon addition of Zn^{2+} ions **(Fig. S2D)**. Increasing the pD up to 6.00 results in progressive deprotonation of histidine imidazoles, thus both peptides bind the zinc ions **(Fig. S2A,B)**. Moving pD towards 6.40 results in precipitation of both peptides (no signal left with 1D ^{1}H NMR spectra) - most likely due to forming larger zinc-coordinated aggregates, oligomers or by simultaneous precipitation of the zinc hydroxide $Zn(OH)_2$, accompanied by precipitation of peptides containing numerous deprotonated hydrophobic imidazole moieties. Interestingly the ATN1mut stays soluble in solution up to pD ca. 0.30 unit higher than the wild-type ATN1.

The "synchronization" effect of the histidine side-chains within the ATN1 can only be explained by the crowding effect of the $(HX)_8$ motif, that "synchronizes" all histidine side-chains to behave similarly in the pH-dependent manner. This effect increases the imidazole side chain dynamics significantly decreasing the ability to bind metal ions, as demonstrated by overlapped NMR signal positions of the imidazole $H\delta2/C\delta2$ and $H\varepsilon1/C\varepsilon1$ correlations **(Fig.**

2F) and zinc titration **(Fig. 2G)**, respectively. Once one of the central histidines from the $(HX)_8$ motif is mutated to tyrosine, all remaining histidines display the behavior more typical to the isolated, not crowded/synchronized histidines, thus the ATN1mut peptide is able to bind zinc ions in an expected mode(s) and pH range. This makes the entire ATN1mut sequence more soluble in the presence of metal ions and at higher pD values closer to the physiological ones.

Supplemental References

1. Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med 369, 1502-1511.
2. Alazami, A.M., Patel, N., Shamseldin, H.E., Anazi, S., Al-Dosari, M.S., Alzahrani, F., Hijazi, H., Alshammari, M., Aldahmesh, M.a., Salih, M.a., et al. (2014). Accelerating Novel Candidate Gene Discovery in Neurogenetic Disorders via Whole-Exome Sequencing of Prescreened Multiplex Consanguineous Families. Cell reports, 1-14.
3. Moutton, S., Bruel, A.L., Assoum, M., Chevarin, M., Sarrazin, E., Goizet, C., Guerrot, A.M., Charollais, A., Charles, P., Heron, D., et al. (2018). Truncating variants of the DLG4 gene are responsible for intellectual disability with marfanoid features. Clin Genet 93, 1172-1178.
4. Tanaka, A.J., Cho, M.T., Millan, F., Juusola, J., Retterer, K., Joshi, C., Niyazov, D., Garnica, A., Gratz, E., Deardorff, M., et al. (2015). Mutations in SPATA5 Are Associated with Microcephaly, Intellectual Disability, Seizures, and Hearing Loss. Am J Hum Genet 97, 457-464.
5. Palmer, E.E., Kumar, R., Gordon, C.T., Shaw, M., Hubert, L., Carroll, R., Rio, M., Murray, L., Leffler, M., Dudding-Byth, T., et al. (2017). A Recurrent De Novo Nonsense Variant in ZSWIM6 Results in Severe Intellectual Disability without Frontonasal or Limb Malformations. Am J Hum Genet 101, 995-1005.
6. Gayevskiy, V., Roscioli, T., Dinger, M.E., and Cowley, M.J. (2018). Seave: a comprehensive web platform for storing and interrogating human genomic variation. Bioinformatics, bty540-bty540.
7. Zimmermann, L., Stephens, A., Nam, S.Z., Rau, D., Kubler, J., Lozajic, M., Gabler, F., Soding, J., Lupas, A.N., and Alva, V. (2018). A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol 430, 2237-2243.
8. Covington, A.K., Paabo, M., Robinson, R.A., and Bates, R.G. (2002). Use of the glass electrode in deuterium oxide and the relation between the standardized pD (paD) scale and the operational pH in heavy water.
9. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-293.
10. Wüthrich, K. (1986). NMR of Proteins and Nucleic Acids.(New York: Wiley).
11. Remelli, M., Brasili, D., Guerrini, R., Pontecchiani, F., Potocki, S., Rowinska-Zyrek, M., Watly, J., and Kozlowski, H. (2018). Zn(II) and Ni(II) complexes with poly-histidyl peptides derived from a snake venom. Inorganica Chimica Acta 472, 149-156.
12. Witkowska, D., Rowinska-Zyrek, M., Valensin, G., and Kozlowski, H. (2012). Specific poly-histidyl and poly-cysteil protein sites involved in $Ni2^+$ homeostasis in H*elicobacter pylori.* Impact of $Bi3^+$ ions on $Ni2^+$ binding to proteins. Structural and thermodynamic aspects. Coordination Chemistry Reviews 256, 133-148.