Automated Fast-Flow Synthesis of *C9orf72* Dipeptide Repeat Proteins

Kohei Sato,[†,‡] Charlotte E. Farquhar,[‡] Bradley L. Pentelute*[,‡,¶,§,#]

[†]School of Life Science and Technology, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama-shi, Kanagawa 226-8501, Japan; [‡]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States; [¶]Center for Environmental Health Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States; [§]The Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, United States; [#]Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, United States.

ABSTRACT: An expansion of the hexanucleotide (GGGGCC) repeat sequence in the chromosome 9 open frame 72 (*c9orf72*) is the most common genetic mutation in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). The mutation leads to the production of toxic dipeptide repeat proteins (DPRs) that induce neurodegeneration. However, the fundamental physicochemical properties of DPRs remain largely unknown due to their limited availability. Here, we synthesized the *c9orf72* DPRs poly-glycine-arginine (poly-GR), poly-proline-arginine (poly-PR), poly-glycine-proline (poly-GP), poly-proline-alanine (poly-PA), and poly-glycine-alanine (poly-GA) using automated fast-flow peptide synthesis (AFPS) and achieved single-domain chemical synthesis of proteins with up to 200 amino acids. Circular dichroism spectroscopy of the synthetic DPRs revealed that proline-containing poly-PR, poly-GP, and poly-PA could adopt polyproline II-like helical secondary structures. In addition, structural analysis by size-exclusion chromatography indicated that longer poly-GP and poly-PA might aggregate. Furthermore, cell viability assay showed that human neuroblastoma cells cultured with poly-GR and poly-PR with longer repeat length resulted in reduced cell viability, while poly-GP and poly-PA did not, thereby reproducing the cytotoxic property of endogenous DPRs. This research demonstrated the potential of AFPS to synthesize low-complexity peptides and proteins necessary for studying their pathogenic mechanisms and constructing disease models.

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are refractory neurodegenerative diseases in which ALS causes progressive muscle paralysis, and FTD causes personality and behavioral changes.[1] These two diseases share many pathologic, genetic, and clinical features, where half of ALS patients are estimated to develop the aspects of FTD.[2] The most common genetic cause of ALS and FTD is a mutation in the chromosome 9 open frame 72 (*c9orf72*) gene with the expansion of the hexanucleotide (GGGGCC) repeat sequence.[3,4] The mutated sequence is further translated by the non-canonical mechanism to produce five dipeptide repeat proteins (DPRs): poly-glycine-arginine (poly-GR), poly-proline-arginine (poly-PR), poly-glycine-proline (poly-GP), poly-proline-alanine (poly-PA), and poly-glycine-alanine (poly-GA).[5,6]

Since the discovery of DPRs, their neurogenerative properties have been extensively studied using DPR-expressing disease models.[7–11] However, the fundamental physicochemical properties of DPRs remain largely unknown due to their limited availability. Although recombinant expression has been the most powerful approach to obtaining proteins, these techniques are limited by scale. In addition, the expression of toxic peptides and proteins, such as highly cationic ones, is often challenging.[12,13] Therefore, a synthetic approach to obtain such DPRs is of great importance to further understand the pathogenic mechanisms of ALS/FTD and to develop new therapies targeting DPRs. In particular, the chemical synthesis of DPRs longer than 36 repeats (72 amino acids) is crucial, as this has been described as the minimum length to exhibit toxicity *in vivo*.[8,11] Although several groups

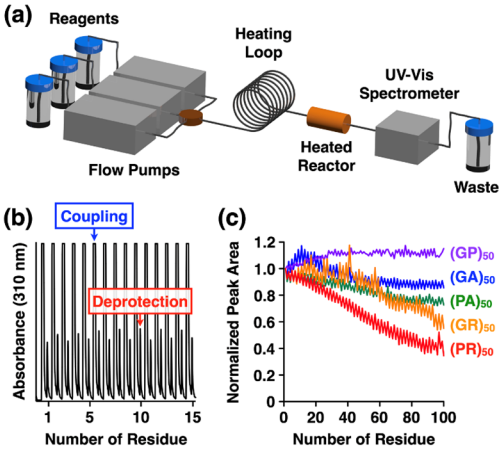


Figure 1. Rapid flow synthesis of dipeptide repeat proteins. (a) Schematic illustration of the automated fast-flow peptide synthesizer (AFPS) used in this study. (b) Absorbance recorded at 310 nm during the synthesis of poly-GP. (c) Normalized peak area of Fmoc deprotections obtained during the synthesis of five DPRs with 50 repeats (100 amino acids).

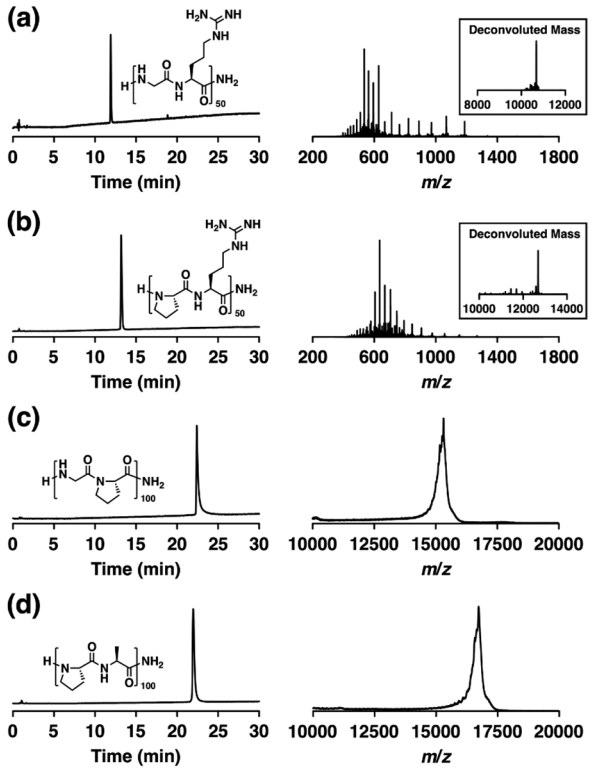


Figure 2. Analytical RP-HPLC traces and mass spectra of (a) **(GR)$_{50}$**, (b) **(PR)$_{50}$**, (c) **(GP)$_{100}$**, and (d) **(PA)$_{100}$** (**(GR)$_{50}$** and **(PR)$_{50}$**: ESI-TOF mass spectrometry, **(GP)$_{100}$** and **(PA)$_{100}$**: MALDI-TOF mass spectrometry).

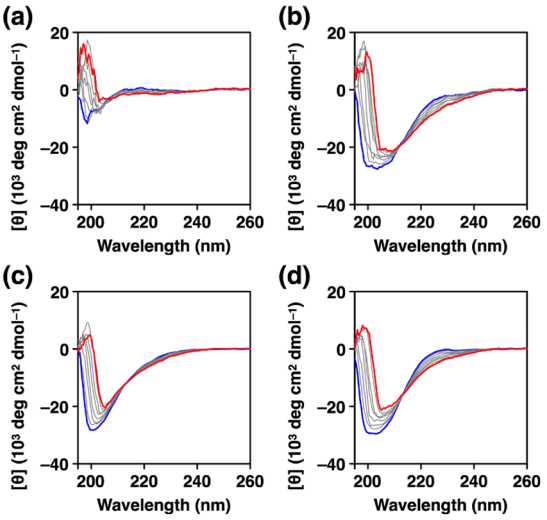


Figure 3. CD spectra of (a) **(GR)$_{50}$**, (b) **(PR)$_{50}$**, (c) **(GP)$_{100}$**, and (d) **(PA)$_{100}$** in PBS ([DPR] = 0.1 mg mL^{-1}) upon heating from 25 °C (blue) to 95 °C (red). Spectra were recorded at every 10 °C.

have reported the chemical synthesis of DPRs, their maximum length is limited to up to 30 repeats (60 amino acids),[14–16] which is reasonable considering the general difficulty of synthesizing peptides longer than 50 amino acids by standard solid-phase peptide synthesis (SPPS).[17] Additionally, chemoselective ligation reactions would not be readily applicable to the synthesis of long DPRs due to the lack of reactive residues.[18] We thus set the goal of this work to overcome these problems and to establish the chemical synthesis of long DPRs to investigate their properties.

Our research group has recently developed an automated fast-flow peptide synthesis (AFPS) system that can complete a single amide coupling reaction in 40 seconds and allow the direct production of peptide chains up to 164 amino acids in hours (Figure 1a).[19–22] Here, we performed AFPS of DPRs on a Rink amide resin following the previously optimized synthetic conditions (see Supporting Information).[20] As activation reagents, we used hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU) for coupling glycine and proline, and (7-azabenzotriazol-1-yloxy)trispyrrolidinophosphonium hexafluorophosphate (PyAOP) for alanine and arginine with extended coupling time. Removal of fluorenylmethylmethoxy-carbonyl (Fmoc) groups was performed using piperidine, and the deprotection reaction was monitored by in-line ultraviolet-visible detection of the eluent (Figure 1b).

Figure 1c shows the normalized area of the deprotection peaks during the synthesis of five DPRs with 50 repeats (100 amino acids). We found that the arginine-containing DPRs **(GR)$_{50}$** and **(PR)$_{50}$** showed relatively smaller peak areas compared to the DPRs variants without arginine **(GP)$_{50}$**, **(GA)$_{50}$**, and **(PA)$_{50}$**, indicating the low coupling efficiency of arginine residues. This observation is consistent with our previous machine learning study showing that arginine residues can have reduced coupling efficiency, probably due to their bulky aromatic 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) protective groups.[23] Nevertheless, AFPS followed by purification using reverse-phase high-performance liquid chromatography (RP-HPLC) allowed the isolation of **(GR)$_{50}$** and **(PR)$_{50}$** in 2% (7.2 mg) and 1% (4.4 mg) yields, respectively, which were still sufficient for the subsequent physicochemical and biological studies. The purity of **(GR)$_{50}$**, **(PR)$_{50}$**, and their shorter repeats was confirmed by analytical RP-HPLC and electrospray-ionization time-of-flight (ESI-TOF) mass spectrometry (Figures 2a, 2b, S1, and S2). For poly-GP and poly-PA, we were even able to synthesize and isolate **(GP)$_{100}$** and **(PA)$_{100}$** in 3% (11.3 mg) and 5% (23.7 mg) yield, respectively. Since poly-GP and poly-PA were poorly ionizable by ESI, we used matrix-assisted laser desorption ionization (MALDI)-TOF mass spectrometry and analytical RP-HPLC to confirm their purity (Figures 2c, 2d, S3, and S4). We also attempted to synthesize poly-GA and isolated **(GA)$_{10}$** and **(GA)$_{20}$** (Figures S5a and S5b). However, we were unable to purify **(GA)$_{50}$** and **(GA)$_{100}$** due to their poor solubility. However, MALDI-TOF mass spectra of crude samples clearly showed the target mass peaks, indicating that AFPS of poly-GA itself was also successful (Figures S5c and S5d).

With the chemically synthesized DPRs in hand, we then performed circular dichroism (CD) spectroscopy of soluble DPRs (poly-GR, poly-PR, poly-GP, and poly-PA) in phosphate-

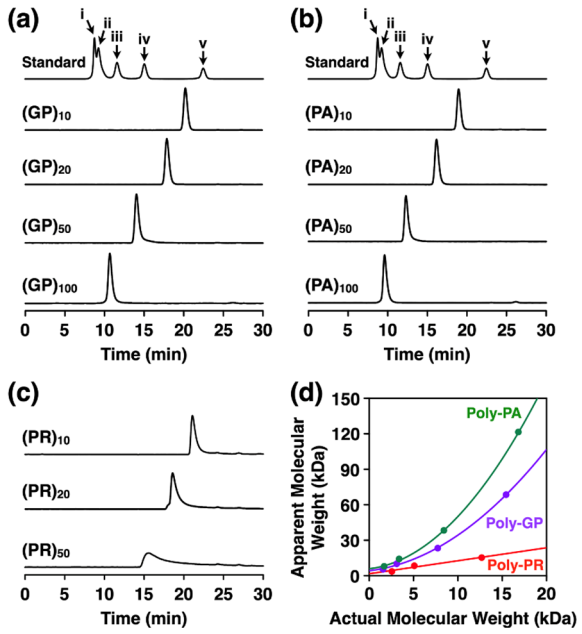


Figure 4. SEC traces of (a) poly-GPs, (b) poly-PAs, and (c) poly-PRs at 25 °C ([DPR] = 1.0 mg mL^{-1} in PBS, eluent: PBS). The molecular weights of the globular protein standards are (i) 670 kDa, (ii) 158 kDa, (iii) 44 kDa, (iv) 17 kDa, and (v) 1.35 kDa, respectively. (d) Plots and their corresponding fitted curves of the actual molecular weight of DPRs as a function of their derived apparent molecular weight using protein standards.

buffered saline (PBS) at different temperatures to investigate their secondary structures. Previously, the secondary structure of poly-GA has been extensively studied and its propensity to form fibrous aggregates through the formation of β-sheet structures has been reported.[14,24–26] However, the structures of other DPRs have not been studied in detail. As shown in Figure 3a, **$(GR)_{50}$** exhibited featureless CD spectra regardless of the measurement temperature, indicating its random coil-rich structure, consistent with previous reports.[10,24] We then measured CD spectra of **$(PR)_{50}$**, **$(GP)_{100}$**, and **$(PA)_{100}$** at 25 °C (Figures 3b–3d, blue curves) and observed a negative peak at 200 nm, which has been previously characterized as flexible or random coil configurations.[10,24,26] However, as we increased the temperature, the peak top showed a red shift associated with the decrease in intensity around 220–230 nm, with an apparent isodichroic point at 215 nm. These results demonstrated that the proline-containing DPRs **$(PR)_{50}$**, **$(GP)_{100}$**, and **$(PA)_{100}$** do not simply adopt disordered structures but show a transition from one state to another upon heating.[27,28] We then found that the observed thermal transition of **$(PR)_{50}$**, **$(GP)_{100}$**, and **$(PA)_{100}$** closely resembles the denaturation profile of collagens.[29] It is noteworthy that approximately 1/3 of the constituent amino acids of collagens are prolines or hydroxyprolines, which contribute to their formation of polyproline II-like helical secondary structures.[29,30] Thus, it is reasonable to assume that **$(PR)_{50}$**, **$(GP)_{100}$**, and **$(PA)_{100}$**, whose half of the constituent amino acids are prolines, also adopt similar secondary structures. In addition, we measured the CD spectra of proline-containing DPRs with shorter repeat lengths and found that they exhibited substantially the same spectral profiles (Figures S7–S9). These results suggest that the secondary structures formed within poly-PR, poly-GP, and poly-PA might be present partially rather than entirely throughout the chains.

In addition to the secondary structures, previous studies have also shown that DPRs can self-assemble into aggregates under physiological conditions.[24,26] However, a comprehensive understanding has been lacking due to the unavailability of DPRs with variable repeat lengths. Here, we took advantage of AFPS, which allowed us to obtain the libraries of DPRs, and systematically investigated their aggregation property by size exclusion chromatography (SEC) using PBS as an eluent. As expected, poly-GP, poly-PA, and poly-PR showed shorter retention times as their molecular weight increased (Figures 4a–4c). However, the SEC profiles of poly-GR did not show sharp peaks under the same analytical conditions (Figure S10a), likely due to their strong interaction with the SEC matrix (agarose and dextran), as indicated by the peak sharpening under the high-salt conditions (Figure S10b). We then evaluated the apparent molecular weight of poly-GP, poly-PA, and poly-PR based on the retention times of globular protein standards. As shown in Figure 4d (red), poly-PR showed a linear correlation between the derived apparent molecular weight and the actual molecular weight (R^2 = 0.97), indicating that poly-PR exists as a monomeric protein and adopts a globular shape. Interestingly, the apparent molecular weight of poly-GP and poly-PA showed an exponential increase (R^2 = 0.99, Figure 4d purple and green). Considering that CD spectral studies of poly-GP and poly-PA did not show any obvious length-dependent changes (Figures S8 and S9), the observed non-linear increase in the apparent molecular weight should not be due to the differences in protein folding but most likely to the enhanced aggregation property of longer DPRs through multivalent interactions.[31,32]

Finally, we investigated the cytotoxicity that synthetic DPRs might possess. Following the previously reported protocol,[14] we performed the MTT cell viability assay using human neuroblastoma cells in the presence of synthesized poly-DPRs. We pre-dissolved synthetic DPRs with different repeat length and concentration in cell media, added the solution of DPRs to BE(2)-C cells and incubated for 48 h, and then quantified the cell viability relative to untreated cells (see the Supporting Information for calculations). As shown in Figures 5a and 5b, poly-GR and poly-PR exhibited lower cell viability with increasing repeat length and concentration. In contrast, poly-GP and poly-PA did not show significant changes in cell viability regardless of their repeat length and concentration (Figures 5c and 5d). This tendency is consistent with previous reports on endogenous DPRs expressed *in vivo*,[8,9] indicating that the synthetic DPRs successfully reproduced their intrinsic cytotoxicity.

In conclusion, we have successfully synthesized *c9orf72* DPRs with up to 100 repeats (200 amino acids) using AFPS. CD spectroscopy of synthetic DPRs revealed that poly-PR, poly-GP, and poly-PA might adopt polyproline II-like helical secondary structures. Furthermore, SEC analysis of synthetic DPRs

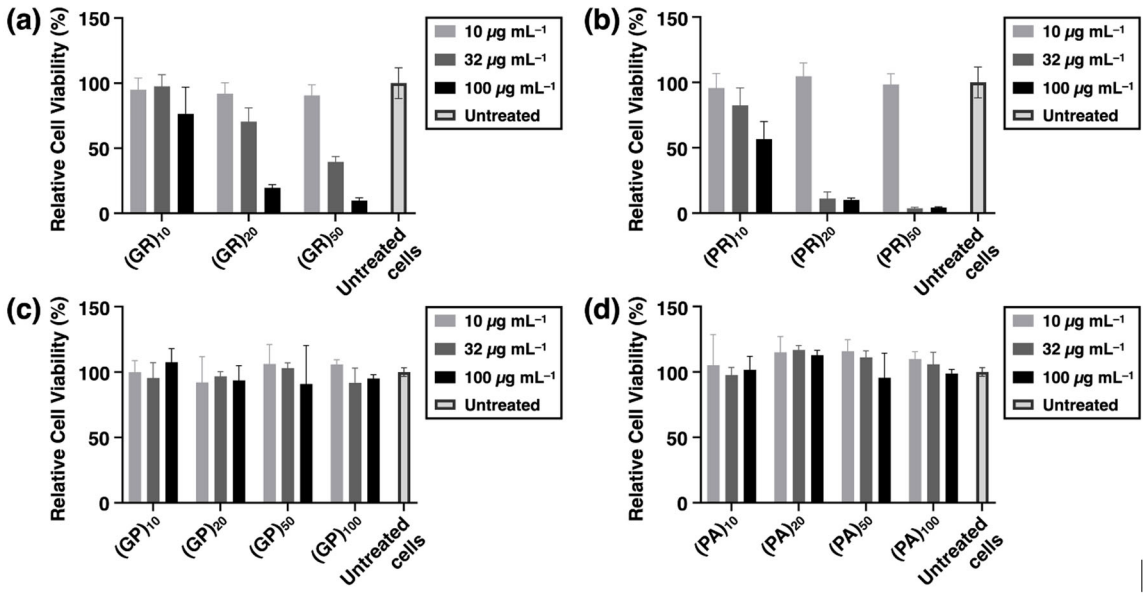


Figure 5. Relative viability of human neuroblastoma BE(2)-C cells 48 h after treatment with media containing (a) poly-GR, (b) poly-PR, (c) poly-GP, and (d) poly-PA with different repeat length and concentrations. The cell viability was quantified by MTT assay as percentage viable relative to untreated cells. Bars represent the standard deviation of the mean based on n=2 replicates.

showed that poly-GP and poly-PA form larger aggregates as their repeat length increases. Finally, a cell viability assay using human neuroblastoma cells demonstrated the increased cytotoxicity of longer poly-GR and poly-PR. We believe that this research has demonstrated the potential of AFPS to synthesize peptides and proteins with limited availability, which is necessary to study their pathogenic mechanisms in detail and to construct disease models for drug discovery.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website. Details of synthesis, characterization, and experimental procedures (PDF).

AUTHOR INFORMATION

Corresponding Author

blp@mit.edu

Notes

B.L.P. is a co-founder and/or member of the scientific advisory board of several companies focusing on the development of protein and peptide therapeutics. The other authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors thank the Department of Chemistry Instrumentation Facility (DCIF) and the Biophysical Instrumentation Facility (BIF) at Massachusetts Institute of Technology for MALDI-TOF mass spectrometry and circular dichroism spectroscopy, respectively. This work was supported by Grant-in-Aid for Early-Career Scientists (21K14670 to K.S.) and Grant-in-Aid for Transformative Research Areas “Molecular Cybernetics” (21H05872 to K.S.).

REFERENCES

(1) Lattante, S.; Ciura, S.; Rouleau, G. A.; Kabashi, E. Defining the genetic connection linking amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD). *Trends Genet.* **2015**, *31*, 263–273.

(2) Ringholz, G. M.; Appel, S. H.; Bradshaw, M.; Cooke, N. A.; Mosnik, D. M.; Schulz, P. E. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* **2005**, *65*, 586–590.

(3) DeJesus-Hernandez, M.; Mackenzie, I. R.; Boeve, B. F.; Boxer, A. L.; Baker, M.; Rutherford, N. J.; Nicholson, A. M.; Finch, N. A.; Flynn, H.; Adamson, J.; Kouri, N.; Wojtas, A.; Sengdy, P.; Hsiung, G.-Y. R.; Karydas, A.; Seeley, W. W.; Josephs, K. A.; Coppola, G.; Geschwind, D. H.; Wszolek, Z. K.; Feldman, H.; Knopman, D. S.; Petersen, R. C.; Miller, B. L.; Dickson, D. W.; Boylan, K. B.; Graff-Radford, N. R.; Rademakers, R. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of *C9ORF72* Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **2011**, *72*, 245–256.

(4) Renton, A. E.; Majounie, E.; Waite, A.; Simón-Sánchez, J.; Rollinson, S.; Gibbs, J. R.; Schymick, J. C.; Laaksovirta, H.; van Swieten, J. C.; Myllykangas, L.; Kalimo, H.; Paetau, A.; Abramzon, Y.; Remes, A. M.; Kaganovich, A.; Scholz, S. W.; Duckworth, J.; Ding, J.; Harmer, D. W.; Hernandez, D. G.; Johnson, J. O.; Mok, K.; Ryten, M.; Trabzuni, D.; Guerreiro, R. J.; Orrell, R. W.; Neal, J.; Murray, A.; Pearson, J.; Jansen, I. E.; Sondervan, D.; Seelaar, H.; Blake, D.; Young, K.; Halliwell, N.; Callister, J. B.; Toulson, G.; Richardson, A.; Gerhard, A.; Snowden, J.; Mann, D.; Neary, D.; Nalls, M. A.; Peuralinna, T.; Jansson, L.; Isoviita, V.-M.; Kaivorinne, A.-L.; Hölttä-Vuori, M.; Ikonen, E.; Sulkava, R.; Benatar, M.; Wuu, J.; Chiò, A.; Restagno, G.; Borghero, G.; Sabatelli, M.; ITALSGEN Consortium; Heckerman, D.; Rogaeva, E.; Zinman, L.; Rothstein, J. D.; Sendtner, M.; Drepper, C.; Eichler, E. E.; Alkan, C.; Abdullaev, Z.; Pack, S. D.; Dutra, A.; Pak, E.; Hardy, J.; Singleton, A.; Williams, N. M.; Heutink, P.; Pickering-Brown, S.; Morris, H. R.; Tienari, P. J.; Traynor, B. J. A Hexanucleotide Repeat Expansion in *C9ORF72* Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **2011**, *72*, 257–268.

(5) Mori, K.; Arzberger, T.; Grässer, F. A.; Gijselinck, I.; May, S.; Rentzsch, K.; Weng, S.-M.; Schludi, M. H.; van der Zee, J.; Cruts, M.; Van Broeckhoven, C.; Kremmer, E.; Kretzschmar, H. A.; Haass, C.; Edbauer, D. Bidirectional transcripts of the expanded *C9orf72* hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.* **2013**, *126*, 881–893.

(6) Zu, T.; Liu, Y.; Bañez-Coronel, M.; Reid, T.; Pletnikova, O.; Lewis, J.; Miller, T. M.; Harms, M. B.; Falchook, A. E.; Subramony, S. H.; Ostrow, L. W.; Rothstein, J. D.; Troncoso, J. C.; Ranum, L. P. W. RAN proteins and RNA foci from antisense transcripts in *C9ORF72* ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E4968–E4977.

(7) Wen, X.; Tan, W.; Westergard, T.; Krishnamurthy, K.; Markandaiah, S. S.; Shi, Y.; Lin, S.; Shneider, N. A.; Monaghan, J.; Pandey, U. B.; Pasinelli, P.; Ichida, J. K.; Trotti, D. Antisense Proline-Arginine RAN Dipeptides Linked to *C9ORF72*-ALS/FTD Form Toxic Nuclear Aggregates that Initiate In Vitro and In Vivo Neuronal Death. *Neuron* **2014**, *84*, 1213–1225.

(8) Mizielinska, S.; Grönke, S.; Niccoli, T.; Ridler, C. E.; Clayton, E. L.; Devoy, A.; Moens, T.; Norona, F. E.; Woollacott, I. O. C.; Pietrzyk, J.; Cleverley, K.; Nicoll, A. J.; Pickering-Brown, S.; Dols, J.; Cabecinha, M.; Hendrich, O.; Fratta, P.; Fisher, E. M. C.; Partridge, L.; Isaacs, A. M. *C9orf72* repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* **2014**, *345*, 1192–1194.

(9) Lee, K.-H.; Zhang, P.; Kim, H. J.; Mitrea, D. M.; Sarkar, M.; Freibaum, B. D.; Cika, J.; Coughlin, M.; Messing, J.; Molliex, A.; Maxwell, B. A.; Kim, N. C.; Temirov, J.; Moore, J.; Kolaitis, R.-M.; Shaw, T. I.; Bai, B.; Peng, J.; Kriwacki, R. W.; Taylor, J. P. C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* **2016**, *167*, 774–788.

(10) Freibaum, B. D.; Taylor, J. P. The Role of Dipeptide Repeats in C9ORF72-Related ALS-FTD. *Front. Mol. Neurosci.* **2017**, *10*, 35.

(11) Sharpe, J. L.; Harper, N. S.; Garner, D. R.; West, R. J. H. Modeling *C9orf72*-Related Frontotemporal Dementia and Amyotrophic Lateral Sclerosis in *Drosophila*. *Front. Cell. Neurosci.* **2021**, *15*, 770937.

(12) Chen, Y. Q.; Zhang, S. Q.; Li, B. C.; Qiu, W.; Jiao, B.; Zhang, J.; Diao, Z. Y. Expression of a cytotoxic cationic antibacterial peptide in *Escherichia coli* using two fusion partners. *Protein Expr. Purif.* **2008**, *57*, 303–311.

(13) Rosano, G. L.; Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* **2014**, *5*, 172.

(14) Chang, Y.-J.; Jeng, U.-S.; Chiang, Y.-L.; Hwang, I.-S.; Chen, Y.-R. The Glycine-Alanine Dipeptide Repeat from C9orf72 Hexanucleotide Expansions Forms Toxic Amyloids Possessing Cell-to-Cell Transmission Properties. *J. Biol. Chem.* **2016**, *291*, 4903–4911.

(15) Boeynaems, S.; Bogaert, E.; Kovacs, D.; Konijnenberg, A.; Timmerman, E.; Volkov, A.; Guharoy, M.; De Decker, M.; Jaspers, T.; Ryan, V. H.; Janke, A. M.; Baatsen, P.; Vercruysse, T.; Kolaitis, R.-M.; Daelemans, D.; Taylor, J. P.; Kedersha, N.; Anderson, P.; Impens, F.; Sobott, F.; Schymkowitz, J.; Rousseau, F.; Fawzi, N. L.; Robberecht, W.; Van Damme, P.; Tompa, P.; Van Den Bosch, L. Phase Separation of *C9orf72* Dipeptide Repeats Perturbs Stress Granule Dynamics. *Mol. Cell* **2017**, *65*, 1044–1055.

(16) Babu, M.; Favretto, F.; de Opakua, A. I.; Rankovic, M.; Becker, S.; Zweckstetter, M. Proline/arginine dipeptide repeat polymers derail protein folding in amyotrophic lateral sclerosis. *Nat. Commun.* **2021**, *12*, 3396.

(17) Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338–351.

(18) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J.-C. M.; Melnyk, O. Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chem. Rev*. **2019**, *119*, 7328–7443.

(19) Mijalis, A. J.; Thomas III, D. A.; Simon, M. D.; Adamo, A.; Beaumont, R.; Jensen, K. F.; Pentelute, B. L. A fully automated flow-based approach for accelerated peptide synthesis. *Nat. Chem. Biol.* **2017**, *13*, 464–466.

(20) Hartrampf, N.; Saebi, A.; Poskus, M.; Gates, Z. P.; Callahan, A. J.; Cowfer, A. E.; Hanna, S.; Antilla, S.; Schissel, C. K.; Quartarano, A. J.; Ye, X.; Mijalis, A. J.; Simon, M. D.; Loas, A.; Liu, S.; Jessen, C.; Nielsen, T. E.; Pentelute, B. L. Synthesis of proteins by automated flow chemistry. *Science* **2020**, *368*, 980–987.

(21) Pomplun, S.; Jbara, M.; Schissel, C. K.; Hawken, S. W.; Boija, A.; Li, C.; Klein, I.; Pentelute, B. L. Parallel Automated Flow Synthesis of Covalent Protein Complexes That Can Inhibit MYC-Driven Transcription. *ACS Cent. Sci.* **2021**, *7*, 1408–1418.

(22) Jbara, M.; Pomplun, S.; Schissel, C. K.; Hawken, S. W.; Boija, A.; Klein, I.; Rodriguez, J.; Buchwald, S. L.; Pentelute, B. L. Engineering Bioactive Dimeric Transcription Factor Analogs via Palladium Rebound Reagents. *J. Am. Chem. Soc.* **2021**, *143*, 11788–11798.

(23) Mohapatra, S.; Hartrampf, N.; Poskus, M.; Loas, A.; Gómez-Bombarelli, R.; Pentelute, B. L. Deep Learning for Prediction and Optimization of Fast-Flow Peptide Synthesis. *ACS Cent. Sci.* **2020**, *6*, 2277–2286.

(24) Flores, B. N.; Dulchavsky, M. E.; Krans, A.; Sawaya, M. R.; Paulson, H. L.; Todd, P. K.; Barmada, S. J.; Ivanova, M. I. Distinct *C9orf72*-Associated Dipeptide Repeat Structures Correlate with Neuronal Toxicity. *PLoS ONE* **2016**, *11*, e0165084.

(25) Guo, Q.; Lehmer, C.; Martínez-Sánchez, A.; Rudack, T.; Beck, F.; Hartmann, H.; Pérez-Berlanga, M.; Frottin, F.; Hipp, M. S.; Hartl, F. U.; Edbauer, D.; Baumeister, W.; Fernández-Busnadiego, R. *In Situ* Structure of Neuronal *C9orf72* Poly-GA Aggregates Reveals Proteasome Recruitment. *Cell* **2018**, *172*, 696–705.

(26) Brasseur, L.; Coens, A.; Waeytens, J.; Melki, R.; Bousset, L. Dipeptide repeat derived from C9orf72 hexanucleotide expansions forms amyloids or natively unfolded structures *in vitro*. *Biochem. Biophys. Res. Commun.* **2020**, *526*, 410–416.

(27) Shimada, H.; Caughey, W. S. Dynamic Protein Structures. *J. Biol. Chem.* **1982**, *257*, 11893–11900.

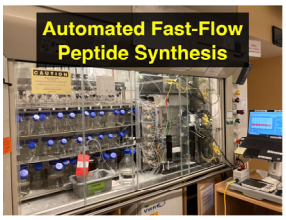
(28) Ryle, M. J.; Lanzilotta, W. N.; Seefeldt, L. C.; Scarrow, R. C.; Jensen, G. M. *J. Biol. Chem.* **1996**, *271*, 1551–1557.

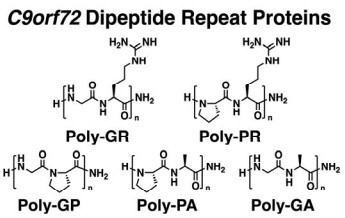
(29) Lopes, J. L. S.; Miles, A. J.; Whitmore, L.; Wallace, B. A. Distinct circular dichroism spectroscopic signatures of polyproline II and unordered secondary structures: Applications in secondary structure analyses. *Protein Sci.* **2014**, *23*, 1765–1772.

(30) Adzhubei, A. A.; Sternberg, M. J. E.; Makarov, A. A. Polyproline-II Helix in Proteins: Structure and Function. *J. Mol. Biol.* **2013**, *425*, 2100–2132.

(31) Mammen, M.; Choi, S.-K.; Whitesides, G. M. Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew. Chem. Int. Ed.* **1998**, *37*, 2754–2794.

(32) Fasting, C.; Schalley, C. A.; Weber, M.; Seitz, O.; Hecht, S.; Koksch, B.; Dernedde, J.; Graf, C.; Knapp, E.-W.; Haag, R. Multivalency as a Chemical Organization and Action Principle. *Angew. Chem. Int. Ed.* **2012**, *51*, 10472–10498.


Automated Fast-Flow
Peptide Synthesis


C9orf72 Dipeptide Repeat Proteins
Poly-GR
Poly-PR
Poly-GP
Poly-PA
Poly-GA