

## Supplementary information

### **Aromatic stacking facilitated self-assembly of ultra-short ionic complementary peptide sequence: $\beta$ -sheet nanofibres with remarkable gelation and interfacial properties**

Jacek K. Wychowaniec,<sup>a,b,c</sup> Ronak Patel,<sup>d</sup> James Leach,<sup>d</sup> Rachel Mathomes,<sup>d</sup> Vikesh Chhabria,<sup>d</sup> Yogita Patil-Sen,<sup>d</sup> Araida Hidalgo-Bastida,<sup>e,f,g</sup> Robert T. Forbes,<sup>d</sup> Joseph M. Hayes,<sup>d</sup> Mohamed A. Elsayy<sup>\*a,b,d,h</sup>

<sup>a</sup> School of Materials, University of Manchester, Oxford Road, Manchester M13 9PL, UK

<sup>b</sup> Manchester Institute of Biotechnology, Oxford Road, Manchester M13 9PL, UK

<sup>c</sup> School of Chemistry, University College Dublin, Belfield, Dublin 4, Ireland

<sup>d</sup> School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK

<sup>e</sup> Centre for Biosciences, Department of Life Science, Manchester Metropolitan University, Manchester, M1 5GD, UK

<sup>f</sup> Centre for Musculoskeletal Science and Sports Medicine, Manchester Metropolitan University, Manchester, M1 5GD, UK

<sup>g</sup> Centre for Advance Materials and Surface Engineering, Manchester Metropolitan University, Manchester, M1 5GD, UK

<sup>h</sup> Leicester Institute of Pharmaceutical Innovation, Leicester School of Pharmacy, De Monfort University, The Gateway, Leicester LE1 9BH, UK

\* Corresponding author: Phone: +44(0)1163664581; e-mail: [mohamed.elsawy@dmu.ac.uk](mailto:mohamed.elsawy@dmu.ac.uk)

## Molecular Dynamics Details:

The initial setup of Phg4 single chain and packed dimer (2×8 peptide chains)  $\beta$ -sheet ladders for molecular dynamics (MD) involved soaking the systems in pre-equilibrated TIP3P model orthorhombic water boxes, allowing for a 10 Å buffer region between peptide and box sides. Details of each system set-up are given in Table S1.

**Table S1.** System set-up details for Phg4 MD simulations

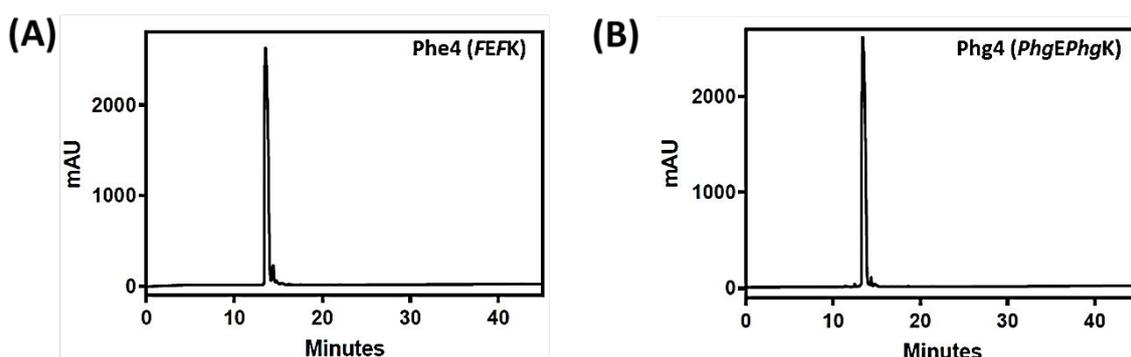
Model	Orthorhombic Box		Number of Water Molecules	Total Number of Atoms
	Dimensions (Å×Å×Å)	Volume (Å <sup>3</sup> )		
Single Chain	32.10×29.78×28.61	27,345	839	2591
Packed Dimer $\beta$ -sheet ladder	43.54×59.85×44.60	116,221	3341	11,207

Forcefield parameters for the peptides were assigned using OPLS3.<sup>1</sup> Heavy-atom bond lengths with hydrogens and the internal geometry of water molecules were constrained using the SHAKE algorithm. Periodic boundary conditions (PBC) were applied and a cut-off of 9.0 Å for non-bond interactions, with electrostatic interactions treated using the Particle Mesh Ewald (PME) method.<sup>2</sup> Before the MD productions runs, a default relaxation protocol for each system was applied: (1) a short Brownian dynamics 100 ps simulation using a temperature (T) of 10 K with solute heavy atoms restrained; (2) a 12 ps simulation in the NVT ensemble using T=10 K (thermostat relaxation constant = 0.1 ps) with solute heavy atoms restrained; (3) a 12 ps simulation in the NPT ensemble using T=10 K (thermostat relaxation constant = 0.1 ps) and pressure (P) =1 atm (barostat relaxation constant = 50 ps) with solute heavy atoms restrained; (4) a 12 ps simulation in the NPT ensemble (T=300 K; thermostat relaxation constant = 0.1 ps; P=1 atm; barostat relaxation constant=50.0 ps) with solute heavy atoms restrained; and (5) a 24 ps simulation in the NPT ensemble (T=300 K; thermostat relaxation constant=0.1 ps; P=1 atm; barostat relaxation constant=2.0 ps) with no restraints. For all the above, the relaxation constants relate to Berendsen thermostats and barostats. After the relaxation, for every model system (Table S1) a 100 ns production run in the NPT ensemble (T=300 K, thermostat relaxation time=1.0 ps; P=1 atm; barostat relaxation time=2.0 ps) was performed using a Nose–Hoover thermostat and Martyna–Tobias–Klein barostat.<sup>3,4</sup> A multiple time step RESPA integration algorithm was used with time steps of 2, 2, and 6 fs for bonded, “near” non-bonded, and “far” non-bonded interactions, respectively. Energy data were recorded every 1.2 ps and atomic coordinate data (1000 frames) from each simulation saved every 100 ps for further analysis. Analysis and visualization of the MD trajectories was performed using Desmond Maestro’s simulation analysis tools. Exploiting Maestro’s Desmond trajectory clustering tool and the affinity propagation method<sup>5</sup>, clustering of trajectory frames (frequency 5, last 80 ns) into clusters was performed based on atomic RMSDs (peptide heavy atoms). The MD simulations were repeated three times for the packed dimer (three independent 100 ns simulations), using the same starting model from the initial Prime sidechain predictions but using different seeds for assignment of the initial velocities. Additionally, Prime v4.5<sup>6</sup> energy calculations were performed on the trajectory frames from the last 80 ns of these simulations with all waters deleted, as a further check of the stability of the simulated peptide aggregation models and for calculation of an aggregation (association) energy for each trajectory according to the following equation:

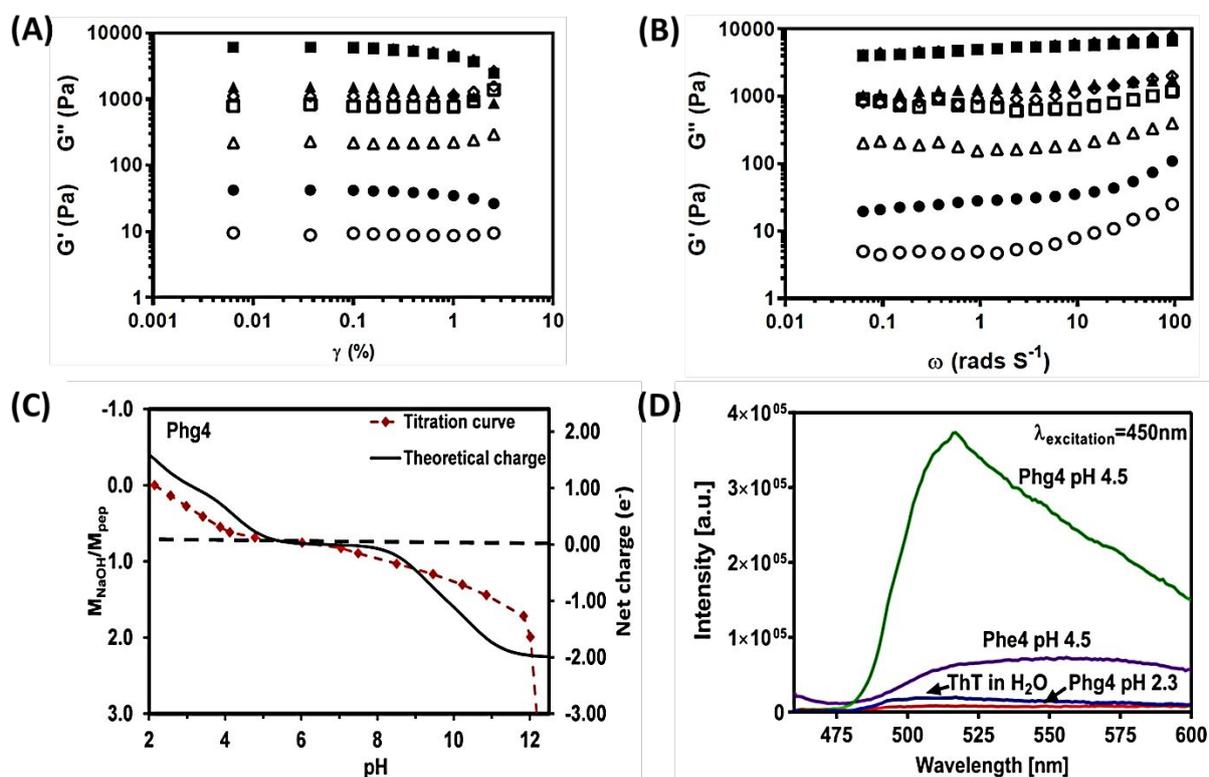
$$\Delta E_{assoc} = \langle V \rangle_{dimer} - (\langle V \rangle_{single\ chain} \times n_c) \quad Eq. S1$$

where  $\langle V \rangle_{dimer}$  is average potential energy from the last 80 ns of each packed dimer model simulation and  $\langle V \rangle_{single\ chain}$  is a constant and the average from the 100 ns single chain MD simulation and  $n_c$  is the number of chains (16 in this case). The OPLS3 forcefield was again employed and water solvation effects accounted for using the implicit variable-dielectric generalized Born model (VSGB).<sup>7</sup> For the semi-empirical PM7 packed dimer optimizations, the last 80 ns of each of the three independent dimer simulations were combined and frames again clustered (frequency 10) using Maestro's Desmond trajectory clustering tool; the representative members from each resulting cluster then used in the PM7 calculations.

### Supporting Figures:



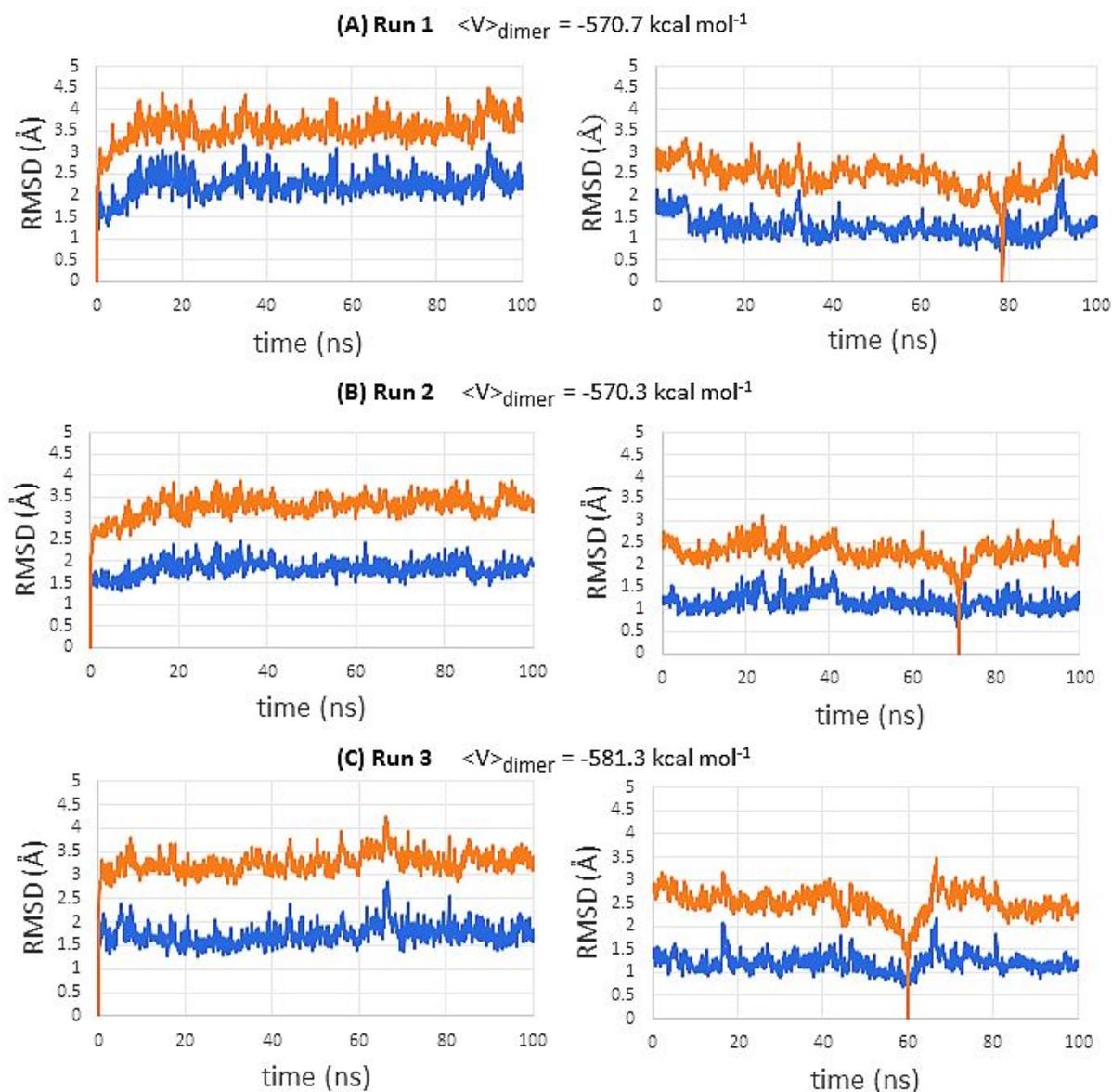
**Figure S1:** RP-HPLC chromatograms of A) Phe4 (FEFK) and B) Phg4 (PhgEPhgK) peptides used. Peptide solutions (1 mg/mL in 1% trifluoroacetic acid in water/acetonitrile 50/50 V/V) injected on Phenomenex Jupiter column (4 $\mu$  Proteo column 90A $^\circ$ , 250x4.66mm) equipped with UV detector ( $\lambda$ 220 nm). Elution gradient was used (flow rate of 1mL/min, started from 90% solvent A (0.05% TFA in H<sub>2</sub>O)/10% solvent B (0.05% TFA in CH<sub>3</sub>CN) to 30% solvent A/70% solvent B in 45 minutes). Peptides purity was estimated as >95% as calculated from area under the peak data.



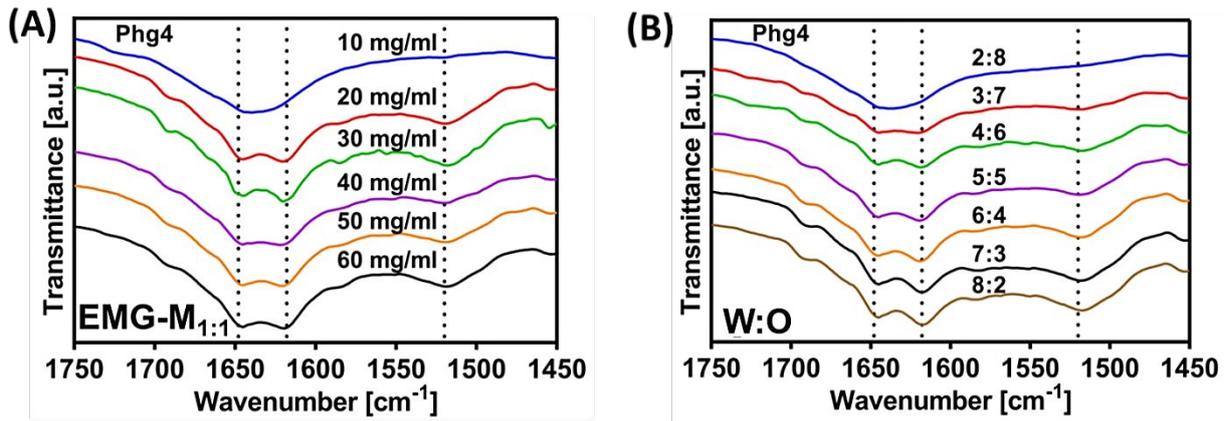
**Figure S2:** Oscillatory rheology characterisation for Phg4 hydrogels A) strain sweep at angular frequency ( $\omega$ ) 6 rad s<sup>-1</sup> and B) Frequency sweep at strain ( $\gamma$ ) 0.2%, within linear viscoelastic region, at peptide concentrations 10 mg mL<sup>-1</sup> (●, ○), 20 mg/mL (▲, △), 30 mg/mL (■, □) and 40 mg/mL (◆, ◇) (close symbols:  $G'$ ; open symbols:  $G''$ ). C) Theoretical net charge state on Phg4 peptide (calculated from equation S1 here below) as a function of pH and molar ratio of added NaOH solution to the Phg4 peptide solution ( $M_{\text{NaOH}}/M_{\text{pep}}$ ) as a function of pH. The experimental titration curve indicates that the net peptide charge is neutral (dashed line indicating zero net charge) within the self-assembly and gelation pH range (pH 4.5-8) when compared to the theoretical charge curve. D) ThT fluorescence showing significant enhancement of intensity for Phg4 at pH 4.5 compared to pH 2.3 indicating the formation of  $\beta$ -sheet fibres, which was marginal for Phe4 at pH 4.5

$$Z = \sum_i N_i \frac{10^{\text{pKa}_i}}{10^{\text{pH}} + 10^{\text{pKa}_i}} - \sum_j N_j \frac{10^{\text{pH}}}{10^{\text{pH}} + 10^{\text{pKa}_j}} \quad \text{Eq. S2}$$

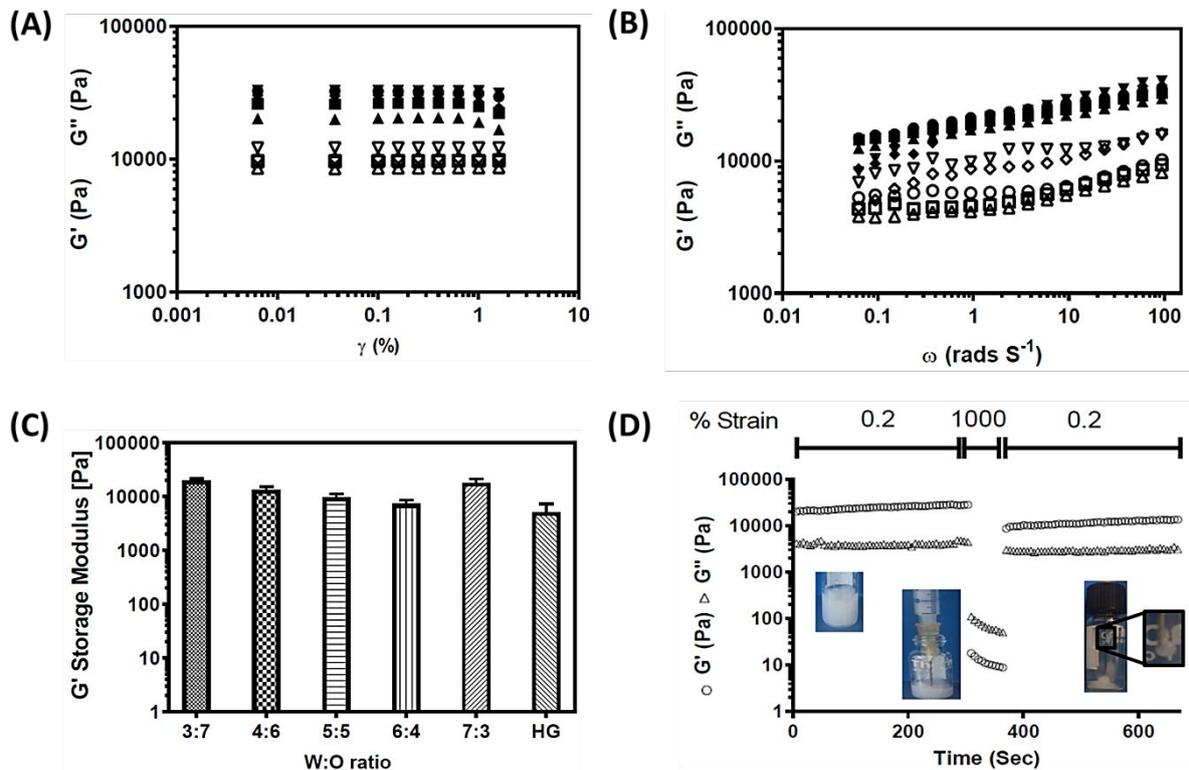
where  $N_i/j$  are the numbers and  $\text{pKa}_i/j$  the  $\text{pKa}$  values of the basic ( $i - \text{pKa} > 7$ ) and acidic ( $j - \text{pKa} < 7$ ) groups present on the peptide.<sup>8</sup>



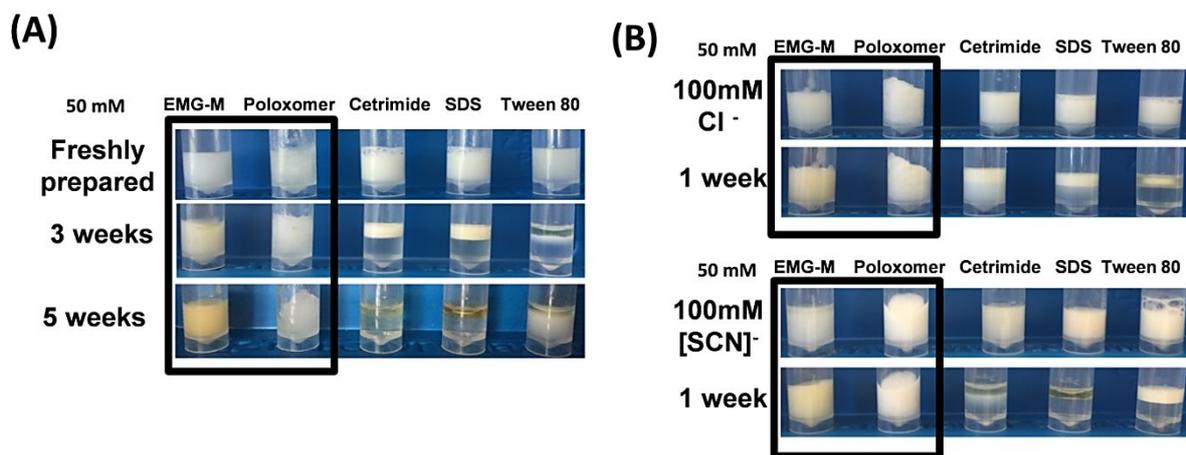
**Figure S3:** RMSDs backbone (blue) and sidechain (orange) of the aggregated packed Phg4 dimer model over the course of the three independent 100 ns MD simulations. On the left, the frame at time 0 ns was used as reference, while on the right, the reference frame was the representative from the largest trajectory cluster. The RMSD between the reference configuration and itself is zero by definition. Backbone RMSDs for the right-hand plots are small (close to  $\sim 1$  Å) and can be associated with thermal fluctuations. The average packed dimer potential energies  $\langle V \rangle_{\text{dimer}}$  are also reported, calculated using Prime and frames from the last 80 ns of each simulation, as described in the Molecular Dynamics Details.



**Figure S4:** A) ATR-FTIR spectra for the of Phg4 Melissa oil mixtures at a series of peptide concentrations (10-60 mg/mL) at 1:1 W:O volume ratio (EMG-M<sub>1:1</sub>), corresponding to the emulgel formation test (Figure 5A). B) ATR-FTIR spectra for EMG-M (30 mg/mL) mixed at a range of different W:O volume ratios (2:8-8:2) corresponding to the emulgel formation test (Figure 5B).



**Figure S5:** Oscillatory rheology characterisation for EMG-M<sub>7:3</sub> emulgels A) strain sweep at angular frequency ( $\omega$ ) 6 rad s<sup>-1</sup> and B) Frequency sweep at strain ( $\gamma$ ) 0.2%, within linear viscoelastic region, at a range of peptide concentrations: 20 mg/mL ( $\blacktriangle$ ,  $\triangle$ ), 30 mg/mL ( $\blacksquare$ ,  $\square$ ), 40 mg/mL ( $\blacklozenge$ ,  $\lozenge$ ), 50 mg/mL ( $\bullet$ ,  $\circ$ ) and 60 mg/mL ( $\blacktriangledown$ ,  $\triangledown$ ) (close symbols: G'; open symbols: G''). C) Shear moduli (G') of EMG-M at a range of different W:O volume ratios (2:8-8:2) compared to hydrogel (HG) at same peptide concentration (30 mg/mL) and pH (5), performed at 6 rad s<sup>-1</sup> obtained from the frequency sweep experiments performed at 0.2% strain D) Strain sweep profiles of EMG-M<sub>7:3</sub> (40 mg/mL) exposed to 0.2% strain for 5 minutes, followed by 1000% strain for 1 minute (simulating injection strain), followed by 0.2% strain for 5 minutes to allow for the gel recovery (O G',  $\Delta$  G'').



**Figure S6:** Emulsion stability profiles of EMG-M<sub>7,3</sub> versus poloxamer (Pluronic F-68), cetrimide (alkyltrimethylammonium bromide), sodium dodecyl sulphate (SDS) and tween 80 at 50mM emulsifier concentration A) storage at room temperature for 3-5 weeks, B) after incubation with 100mM sodium chloride (top panel) and potassium thiocyanate (bottom panel) for 1 week.

### Supporting Video:

Molecular Dynamics video for Phg4 packed dimer of the simulation with  $\Delta E_{assoc} = -550.9$  kcal mol<sup>-1</sup> is attached to the ESI.

### References

- (1) Harder, E.; Damm, W.; Maple, J. Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12* (1), 281–296.
- (2) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593.
- (3) Martyna, G. J.; Klein, M. L. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97* (4), 2635–2643.
- (4) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101* (5), 4177–4189.
- (5) Frey, B. J.; Dueck, D. Clustering by passing messages between data points. *Science* **2007**, *315* (5814), 972–976.
- (6) Schrodinger Release 2016-3, Schrodinger LLC, New York, NY, **2016**.
- (7) Li, J.; Abel, R.; Zhu, K.; Cao, Y.; Zhao, S.; Friesner, R. A. The VSGB 2.0 model: a next generation energy model for high resolution protein structure modelling. *Proteins* **2011**, *79* (10), 2794–2812.
- (8) Lehninger A. L. Principles of Biochemistry. Worth Publishers, New York **1982**.