Electronic Supplementary Material (ESI) for ChemComm. This journal is © The Royal Society of Chemistry 2017

Protein Sensing and Discrimination Using High Functionalised Ruthenium (II) tris(Bipyridyl) Protein Surface Mimetics in an Array Format

Sarah H. Hewitt^{a,b} and Andrew J. Wilson^{a,b*}

^aSchool of Chemistry, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK.

^bAstbury Centre for Structural Molecular Biology, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK

E-mail: a.j.wilson@leeds.ac.uk

Supporting Information

General Considerations

Reagents and solvents were purchased from major suppliers and used without further purification. Anhydrous chloroform, dichloromethane, and methanol were obtained from the in-house solvent purification system, from Innovative Technology Inc. PureSolv[®], other solvents used were of HPLC grade. Water for aqueous solutions was deionised.

Thin layer (silica) chromatography was performed using Merck Kiesegel 60 F_{254} 0.25 mm precoated aluminium plates. Product spots were visualised by colour and under UV light (254 nm and 365 nm). Flash column chromatography was performed using silica gel 60 (0.043 – 0.063 mm VWR or Sigma Aldrich) or alumina (Brockman I from Sigma Aldrich), unless otherwise stated silica gel was used and pressure was applied by means of head bellows.

¹H NMR spectra were obtained on Bruker DPX 300 (300 MHz) Avance 500 (500 MHz) or DRX500 (500 MHz) spectrometers and referenced to either residual non-deuterated solvent peaks or tetramethylsilane. ¹³C spectra were recorded on a Bruker DPX 300 (75 MHz) Bruker or an Avance 500 (126 MHz) and referenced to the solvent peak. ¹H spectra are reported as follows: ¹H NMR (spectrometer frequency, solvent) δ ppm to 2 d.p. (multiplicity, *J* coupling constant in Hertz, number of protons, assignments). Chemical shifts are quoted in ppm with signal splitting recorded as singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m), broad (br). Coupling constants, *J*, are measured to the nearest 0.1 Hz. Similarly, ¹³C spectra are reported as follows: δ (spectrometer frequency, solvent): δ ppm to one decimal place. Assignments of spectra were assisted by the results of DEPT, COSY, HMQC and HMBC experiments.

Infrared spectra were recorded on a Perkin Elmer Fourier-Transfer spectrometer. Spectra were analysed neat and structurally important absorptions are quoted. Absorption maxima (v_{max}) are quoted in wavenumbers (cm⁻¹).

HPLC LC/MS were recorded on a Bruker HCT ultra under electrospray ionisation (ESI) conditions. High resolution mass spectra were recorded on a Bruker Daltonics micrOTOF Premier Mass Spectrometer, under positive ESI conditions unless otherwise stated.

Syntheses

Compounds **1**, **3** and **6-8** were synthesized as described previously.^{1, 2} The NOXA peptide used in this work is shown below and represents a variant on the murine NOXA-B wild type sequence – its synthesis of further study will be described elsewhere.

Wild type *m*NOXA-B:

AAQLRRIGDKVNLRQKLLN

Structure of variant *m*NOXA-B used in this work:

```
FITC-(Ga)AAQLARIGDKVNLRQKLLN-NH2
```

Ethyl 2-[(2-{4- [(2-ethoxy-2-oxoethyl) carbamoyl] pyridin-2-yl} pyridin-4-yl) formamido]acetate



2,2' bipyridine 4,4' carboxylic acid,^{1, 2} (1.50 g, 6.14 mmol) and thionyl chloride (20 mL) were heated under reflux for 16 hours, then the solvent removed *in vacuo*, the dry acid chloride was added anhydrous chloroform (40 mL), ethyl glycine hydrochloride salt (1.89 g, 13.5 mmol) and triethylamine (1.88 mL, 13.5 mmol) and the reaction heated under reflux, under a nitrogen atmosphere, for 16 hours. The reaction was then cooled and concentrated to yield the crude product as a pink solid, which was purified by flash column chromatography (5 % methanol in dichloromethane) to yield the product as a beige solid (1.52 g, 3.65 mmol, 60 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.37 (t, *J* = 7.2 Hz, 4 H, H7), 4.24 - 4.40 (m, 8 H, H5 and H6), 7.93 (d, *J* = 4.8 Hz, 2 H, H2), 8.82 (s, 2 H, H1), 8.91 (d, *J* = 4.8 Hz, 2 H, H3); ¹³C NMR (126 MHz, DMSO-*d*₆) δ ppm 14.1, 41.4, 60.6, 118.2, 121.9, 142.1, 150.2, 155.6, 165.1, 169.5; IR (solid state, cm⁻¹) 3303 (N-H), 1741 (C=O ester), 1648 (C=O amide); ESI-MS *m/z* found 415.16203 [M+H]⁺, [C₂₀H₂₃N₄O₆]⁺ requires 415.1612

Tris (ethyl 2-[(2-{4-[(2-ethoxy-2-oxoethyl)carbamoyl]pyridin-2-yl}pyridin-4-yl)formamido]acetate) ruthenium(II) dinitrate



Ethyl 2-[(2-{4-[(2-ethoxy-2oxoethyl) carbamoyl] pyridin-2-yl pyridin-4-yl) formamido]acetate, (410 mg, 0.991 mmol), Ru(DMSO)₄Cl₂ (150 mg, 0.310 mmol) and silver nitrate (105 mg, 0.620 mmol) in ethanol (20 mL) were heated under reflux for 7 days. The resulting solution was cooled to room temperature, filtered and the red filtrate concentrated to yield the crude product as a red solid. This was purified by flash column chromatography (5 % methanol in dichloromethane) to yield the product as a red solid (dinitrate salt, 426 mg, 0.290 mmol, 94 %).¹H NMR (500 MHz, CDCl₃) δ ppm 1.16 (t, J = 6.8 Hz, 18 H, H7), 3.97 -4.18 (m, 48 H, H5 + H6), 7.71 (br. s, 6 H, H2), 9.12 (s, 6 H, H1), 9.34 (s, 6 H, H3); IR (solid state, cm⁻¹) 3256, (N-H), 1734 (C=O ester), 1664 (C=O amide); ESI-MS m/z found m/z 672.1840 [M]²⁺, [C₆₀H₆₆N₁₂O₁₈Ru]²⁺ requires 672.1834

Tris (2- [(2- {4-[(carboxymethyl) carbamoyl] pyridin-2-yl} pyridin-4-yl) formamido]acetic acid) ruthenium(II) dichloride 2



Tris (ethyl 2- [(2-{4- [(2- ethoxy-2- oxoethyl) carbamoyl] pyridin-2- yl} pyridin-4yl)formamido]acetate) ruthenium(II) dinitrate, (200 mg, 0.136 mmol) was dissolved in ethanol (5 mL), and water (5 mL), 1 M sodium hydroxide solution (5 mL) was added, and the resulting mixture stirred for 18 hours. The reaction mixture was then neutralized with 1 M hydrochloric acid and concentrated to yield the product as a red solid in a mixture with sodium chloride. This mixture was dialysed (MWCO 0.5 – 1 kDa) against pure water to yield the product as a dark red solid (173 mg, 0.133 mmol, 98 %); ¹H NMR (500 MHz, D₂O) δ ppm 3.95 (s, 12 H, H5), 7.73 (dd, *J* = 5.9, 1.6 Hz, 6 H, H2), 7.94 (d, *J* = 5.9 Hz, 6 H, H1), 8.99 (s, 6 H, H5); IR (solid state, cm⁻¹) 3251 (O-H), 1644 (acid C=O), 1585 (amide C=O); ESI-MS *m/z* found 588.0885 [M]²⁺, [C₄₈H₄₂N₁₂O₁₈Ru]²⁺ requires 588.0819

Tetraethyl-2,2',2"'-(([2,2'-bipyridine]-4,4'-dicarbonyl)bis(azanetriyl))

tetra-acetate



2,2'-bipyridine-4,4'-dicarboxylic acid (500 mg, 2.05 mmol), triethylamine (1 drop) and thionyl chloride (20 mL) were heated under reflux for 16 hours. The mixture was cooled to room temperature and the thionyl chloride removed *in vacuo* to yield the acid chloride as an orange solid, which was flushed with nitrogen and used immediately. The dry acid chloride was redissolved in anhydrous chloroform (40 mL) and diethyl iminodiacetate (0.80 mL, 4.46 mmol) was added dropwise, followed by anhydrous triethylamine (0.63 mL, 4.52 mmol). The mixture was refluxed for 48 hours. The solution was concentrated and the resulting white solid purified by flash column chromatography (1:9 MeOH:CH₂Cl₂) to yield the product as a yellow solid (1.11 g, 1.89 mmol, 92 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.27 (t, *J* = 7.2 Hz, 6 H, H6/H9), 1.34 (t, *J* = 7.2 Hz, 6 H, H6/H9), 4.11 (s, 4 H, H4/H7), 4.22 (m, 4 H, H5/H8), 4.38 (s, 4 H, H4/H7), 7.40 (dd, *J* = 4.9, 1.5 Hz, 2 H, H1), 8.50 (app. s, 2 H, H3), 8.74 (d, *J* = 4.9 Hz, 2 H, H2); ¹³C NMR (126 MHz, CDCl₃) δ ppm 14.1, 14.2, 47.5, 50.2, 60.9, 61.6, 118.4, 121.5, 143.5, 149.9, 156.0, 168.7, 169.9, 171.7; IR (solid state, cm⁻¹) 1736 (amide C=O); ESI-HRMS found *m/z* 587.2388 [M]⁺, [C₂₈H₃₅N₄O₁₀]⁺ requires 587.2353

Tris(tetraethyl-2,2',2",2"'-(([2,2'-bipyridine]-4,4'-dicarbonyl)bis(azanetriyl))

tetra-acetate ruthenium(II) dinitrate



Tetra-ethyl-2,2',2'',2'''-(([2,2'-bipyridine]-4,4'-dicarbonyl)bis(azanetriyl))tetra-acetate (390 mg, 0.665 mmol), Ru(DMSO)₄Cl₂ (100 mg, 0.206 mmol), silver nitrate (71.0 mg, 0.418 mmol) and ethanol (30 mL) were heated under reflux for 7 days. The reaction mixture was then filtered hot and concentrated to yield the product as a red solid (0.41 g, 0.21 mmol, 100 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.27 (t, *J* = 7.1 Hz, 18 H, H6/H9), 1.34 (t, *J* = 7.2 Hz, 18 H,

H6/H9), 4.25 (q, J = 7.2 Hz, 12 H, H5/H8), 4.23 (q, J = 7.1 Hz, 12 H, H5/H8), 4.36 (br. s., 12 H, H4/H7), 4.35 (br. s., 12 H, H4/H7), 7.65 (dd, J = 5.8, 1.4 Hz, 6 H, H1), 7.96 (d, J = 5.8 Hz, 6 H, H2), 8.55 (s, 6 H, H3); IR (solid state, cm⁻¹) 2982 (CH), 1732 (ester C=O), 1650 (amide C=O) cm⁻¹; ESI-HRMS found *m/z* 930.2967 [M]²⁺, [C₈₄H₁₀₂N₁₂O₃₀Ru]²⁺ requires 930.2931

Tris(2,2',2'',2'''-(([2,2'-bipyridine]-4,4'-dicarbonyl)bis(azanetriyl))tetraacetic acid) ruthenium(II) dichloride



Tris(tetra-ethyl-2,2',2",2"'-(([2,2'-bipyridine]-4,4'-dicarbonyl)bis(azanetriyl))tetra-acetate ruthenium(II) dichloride (50.0 mg, 0.0275 mmol) was dissolved in ethanol (2.5 mL) and water (2.5 mL) and then mixed with sodium hydroxide (2.4 mL, 1M). The mixture was left to stir for 16 hours and then neutralized and concentrated. The salt was removed by dialysis (float-alyser MWCO 0.1–0.5 kDa) against pure water to yield the product as a red solid (42.0 mg, 0.026 mmol, 100 %); ¹H NMR (500 MHz, D₂O) δ ppm 3.88 (s, 12 H, H4/H5), 4.06 (d, *J* = 17.0 Hz, 6 H, H4/H5), 4.11 (d, *J* = 17.0 Hz, 6 H, H4/H5), 7.46 (d, *J* = 5.8 Hz, 6 H, H1), 7.85 (d, *J* = 5.8 Hz, 6 H, H2), 8.61 (s, 6 H, H3); ESI-HRMS found *m/z* 760.0842 [M]²⁺, [C₆₀H₅₄N₁₂O₃₀Ru]²⁺ requires 762.1056; λ_{max} : 298 nm (ϵ / dm³ mol⁻¹ cm⁻¹ 57029)

Tris (*N4, N4'- bis* (6-boc aminohexyl) -2,2'- bipyridine- 4,4'- dicarboxamide) ruthenium(II) dichloride



Tris (2,2'-bipyridine-4,4'-dicarboxylic acid) ruthenium(II) dichloride(114 mg, 0.125 mmol) was heated under reflux in thionyl chloride (30 mL) and dimethylformamide (1 drop) for 6 hours. The solvent was removed *in vacuo* and the resulting red acid chloride flushed with nitrogen and used immediately. The acid chloride was resuspended in anhydrous chloroform (30 mL) and heated to reflux, under a nitrogen atmosphere. *N*-Boc-1,6-diamino hexane (0.25 mL, 1.1 mmol) and anhydrous di*iso*propylethylamine (0.39 mL, 2.3 mmol) were added, and the resulting solution heated under reflux for 16 hours. The reaction mixture was then allowed to

cool to room temperature, and the reaction mixture quenched with saturated sodium hydrogen carbonate solution (30 mL). The aqueous layer was removed and the organic phase washed with 1 M hydrochloric acid (30 mL) and brine (30 mL). The organic phase was dried (sodium sulfate) and concentrated in order to yield the crude product as a red solid. This was purified by flash column chromatography (10 % methanol in dichloromethane) to yield the product as a red solid (95 mg, 0.045 mmol, 36 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.34 (br. d, *J* = 5.2 Hz, 12 H, H7/H8), 1.39 (s, 54 H, H11), 1.44 - 3.47 (br. s, 12 H, H7/H8), 1.53 (m, 12 H, H6/H9), 1.69 (br. s, 24 H, H6/H9 + H5/H10), 3.09 (br. s, 12 H, H5/H10), 4.77 (br. s, 6 H, NHBoc), 7.66 (br. s, 6 H, H2), 8.06 (br. s, 6 H, H1), 8.97 (br. s, 6 H, H3), 10.02 (br. s, 6 H, H4); IR (solid state, cm⁻¹) 3291 (N-H), 1657 (C=O amide); ESI-HRMS found *m/z* 1011.5489 [M]²⁺, [C₁₀₂H₁₅₆N₁₈O₁₈Ru]²⁺ requires 1011.5444

Tris (*N4,N4'-bis* (6-aminohexyl) -2,2'- bipyridine-4,4'- dicarboxamide) ruthenium(II) dichloride



Tris (*N4,N4'-bis*(6-Boc aminohexyl)-2,2'-bipyridine-4,4'-dicarboxamide) ruthenium(II) dichloride (20 mg, 0.0095 mmol) was stirred in 1 M hydrogen chloride in dioxane (5 mL) and water (0.5 mL) for 2 hours. The resulting mixture was concentrated and redissolved in water (10 mL). The solution was neutralised by addition of 1 M sodium hydroxide solution. The neutral solution was concentrated to ~2 mL and the resulting solution dialysed (MWCO 0.1 - 0.5 kDa) against pure water to yield the product as a red solid (14 mg, 0.0094 mmol, 98 %); ¹H NMR (500 MHz, D₂O) δ ppm 1.37 (br. s, 24 H, H7 + H8), 1.62 (br. s, 24 H, H6 + H9), 2.94 (t, *J* = 7.5 Hz, 12 H, H5/H10), 3.38 (t, *J* = 6.8 Hz, 12 H, H5/H10), 7.67 (d, *J* = 5.8 Hz, 6 H, H3), 7.90 (d, *J* = 5.8 Hz, 6 H, H2), 8.89 (s, 6 H, H1); IR (solid state, cm⁻¹) 3386 (N-H), 3255 (N-H), 1717 (C=O amide); ESI-MS *m/z* found 711.3884 [M]²⁺, [C₇₂H₁₀₈N₁₈O₆Ru]²⁺ requires 711.3871

Luminescence Analyses

All arrays were performed in 5 mM sodium phosphate, pH 7.5 buffer. All protein stocks, other than McI-1 and *h*DM2, were made up from freeze-dried protein, purchased from major suppliers, into 5 mM sodium phosphate, pH 7.5 at ~ 1 mM concentration. Accurate concentrations were determined by UV/Vis calculated using extinction coefficients (at 280 nm) for lysozyme, α -chymotrypsin, BSA, papain and ribonuclease A and (at 550 nm) for cyt c, as described elsewhere. All arrays were performed in 384 Optiplate well plates and were scanned using a Perkin Elmer EnVisionTM 2103 MultiLabel plate reader.

<u>*Ru(bpy)*₃ complexes and proteins</u>

To each well was added 20 μ L of Ru(bpy)₃ complex solution 5 μ M and 20 μ L of protein solution 20 μ M or 5 mM sodium phosphate, pH 7.5 buffer. On each plate all wells were measured in triplicate. The plate was incubated for 45 minutes before scanning using fixed wavelengths, excitation 467 nm, emission 630 nm, and using monochromators, excitation 467 nm, emission range 500 – 800 nm, 3 nm step. The peak maxima/ intensities were taken. The values without protein were averaged over the triplicate wells, and the percentage difference for each of the other wells containing the same complex calculated. Each of these results was used for statistical analysis.

Ru(bpy)₃ complexes, FITC-NOXA B (R-A) tracer and proteins

To each well was added 20 μ L of 7.5 μ M Ru(bpy)₃ complex solution (or 5 mM sodium phosphate buffer, pH 7.5), 20 μ L of 30 μ M protein solution (or 5 mM sodium phosphate, pH 7.5 buffer) and 20 μ L of 1.5 μ M FITC-NOXA-B (R-A) peptide (or 5 mM sodium phosphate, pH 7.5 buffer). On each plate wells without protein wells were run in quadruplicate. The plate was incubated for 2 and 20 hours prior to scanning using fixed wavelengths, excitation 467 nm, emission 630 nm, and using monochromators, excitation 467 nm, emission range 480 - 750 nm, 3 nm step. The peak maxima/intensities for both luminescence bands were taken (emission 520 nm and 630 nm). The values obtained for wells with no protein present were averaged over the quadruplicate wells, all other data used was from each well individually. The percentage difference for each well for that with protein to without protein was calculated and used for statistical analysis.

Linear discriminant analysis

Microsoft Excel was used to calculate the percentage differences from no protein for each of the individual wells. Linear discriminant analysis was carried out using XLstat software, then plotted using OriginPro 9.

Below some of the mathematics behind this technique is outlined. Let the vectors **a**, **b**, **c** ... **x** etc. be each of the individual protein replicates

$$\boldsymbol{\alpha} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_8 \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_8 \end{pmatrix}, \ \boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_8 \end{pmatrix} etc.$$

These samples **a**, **b**, **c** etc. also have class labels y_a , y_b , y_c etc.

This means that we can define separate mean vectors for each class (i), μ_i , and we can take the total number of samples in class i, as M_i . Such that-

$$M = \sum_{i=1}^{C} M_i$$

Where C is the total number of classes

From these parameters it is possible to define two scatter matrices, the within class scatter matrix (S_w) and the between class scatter matric (S_B).

$$S_{w} = \sum_{i=1}^{C} \sum_{j=1}^{M_{i}} (y_{j} - \mu_{i})(y_{i} - \mu_{i})^{T}$$
$$S_{B} = \sum_{i=1}^{C} (\mu_{i} - \mu)(\mu_{i} - \mu)^{T}$$

Where **µ** is the mean of the whole data set- $\mu = \frac{1}{c} \sum_{i=1}^{c} \mu_i$

These scatter matrices can be transformed onto new planes (W) using the equations-

$$\tilde{S}_w = W^T S_w W$$

 $\tilde{S}_B = W^T S_B W$

The aim of LDA is to find an optimal W (W^{*}) such that det S_b /det S_w is maximised, i.e. finding the minimum within class scatter to the maximum between class scatter.

$$W^* = argmax \left\{ \frac{W^T S_B W}{W^T S_W W} \right\}$$

This is found by finding a matrix W^* whose columns are the eigenvectors (w_i) corresponding to the largest eigenvalues of the following equation-

$$(S_B - \lambda_i S_w) w_i = 0$$

These eigenvectors and eigenvalues can again be found computationally

Supplementary Data Figures



Figure S1. Differential luminescence responses from the different $Ru(bpy)_3$ surface mimetics **1-8** (2.5 μ M) on incubation with various different proteins (10 μ M), (5 mM sodium phosphate, pH 7.5, $\lambda_{exc} = 467$ nm). a) and b) Illustrative luminescence intensity over variable wavelengths for $Ru(bpy)_3$ complex **2** (a) and **6** (b), graphs obtained using plate reader monochromators.

We tested a range of fluorescently labelled peptides available in our laboratory; of those we tested, the luminescence intensity of Ru(bpy)₃ complexes was found to differ in the presence of the NOXA-B peptide (see e.g. Fig. S4b and d below) in the concentration regime at which the original array experiments were performed hence we used this peptide. Our hypothesis on the sensing mechanism is that in the concentration regime tested, the differential NOXA-B dependent change in Ru(bpy)₃ luminscence intensity is consistent with recognition between the two and upon addition of the protein analyte, this equilibrium is disturbed; the protein may recognise one (competitively) or both fluorescence ligands (either independently or as a ternary complex) to promote a change in fluorescence of both ligands (Fig. S2)



Figure S2. Cartoon illustrating potential protein sensing mechanisms using $Ru(bpy)_3$ complexes and a fluorescently labelled peptide.



Figure S3. Concentration test to determine appropriate concentration of FITC-NOXA B peptide to use with 2.5 μ M Ru(bpy)₃ surface mimetic. Luminescence emission intensity at variable wavelengths (λ_{exc} 467 nm) upon incubation of 2.5 μ M Ru(bpy)₃ surface mimetic **1** with various concentrations of FITC-NOXA B peptide in 5 mM sodium phosphate, pH 7.5



Figure S4. Luminescence response (λ_{exc} = 467 nm, in 5 mM sodium phosphate, pH 7.5, 2 hour incubation) of the FITC-NOXA B peptide alone (a), Ru(bpy)₃ surface mimetic **6** alone (b), FITC-NOXA B with Ru(bpy)₃ surface mimetic **6** (c) and FITC NOXA B with Ru(bpy)₃ surface mimetic **2**, highlighting differential spectral responses with two fluorophores.



Figure S5 2D LDA analyses of Ru(bpy)₃ surface mimetic **1-6** array in the presence of FITC-NOXA B, and in response to six different proteins (a) after 2 hrs incubation (b) after 20 hours incubation (c) Combined of the data after 2 and 20 hrs incubation.(conditions as given in Fig. 2)



Figure S6 3D LDA analyses of Ru(bpy)₃ surface mimetic **1-8** array in the presence of FITC-NOXA B, and in response to eight different proteins (each panel represents a different orientation of the data to aid visualization, conditions as given in Fig. 2)



Figure S7 LDA analyses of Ru(bpy)₃ surface mimetic **1-8** array in the presence of FITC-NOXA B: 95 % confidence ellipsoids for the data shown in Fig. 4a, (each panel represents a different orientation of the data to aid visualization, conditions as given in Fig. 2)

References

- 1. S. H. Hewitt, M. H. Filby, E. Hayes, L. T. Kuhn, A. P. Kalverda, M. E. Webb and A. J. Wilson, *ChemBioChem*, 2017, **18**, 223.
- 2. J. Muldoon, A. E. Ashcroft and A. J. Wilson, *Chem. Eur. J.*, 2010, **16**, 100.