Supporting Information

Selective targeting of disease-relevant protein binding domains by O-phosphorylated natural product derivatives

Martin Gräber, Weronika Janczyk, Bianca Sperl, Nagarajan Elumalai, Christian Kozany,
Felix Hausch, Tad A. Holak, Thorsten Berg*

Contents

1. Supporting Figures

- Figure S1: Activity of Fosfosol is independent of the presence or absence of DTT and the time of incubation with STAT5b
- Figure S2: Re-synthesized fosfosol is equally potent as the commercially available material
- Figure S3: The binding mode of fosfosol (**1a**) to the STAT5b SH2 domain as predicted by molecular modeling is similar as the binding mode of the phosphotyrosine side chain of a peptide that binds to the STAT1 SH2 domain as determined by X-ray crystallography
- Figure S4: Activity of dexamethasone-21-phosphate (**2a**) is independent of the presence or absence of DTT and the time of incubation with Pin1
- Figure S5: Chemical structures and activities of dexamethasone (**2c**) and dexamethasone-21-acetate (**2d**)
- Figure S6: Betamethasone-21-phosphate (**2b**) displays only weak activity against the Pin1 substrate binding domain
- Figure S7: Dexamethasone-21-phosphate (**2a**) binds to the substrate binding domain with significantly higher affinity than to the WW domain
- Figure S8: ¹H, ¹⁵N-HSQC-NMR of full-length Pin1 and dexamethasone-21-phosphate (**2a**)
- Figure S9: ¹H, ¹⁵N-HSQC-NMR of full-length Pin1 and betamethasone-21-phosphate (**2b**)
- Figure S10: Alternative depictions of the X-ray structure of the complex of Pin1 (Arg14Ala mutant) and dexamethasone-21-phosphate (**2a**).
- Figure S11: Chemical structures and chemical names of the 12 O-phosphorylated natural products or natural product derivatives within the 5019-member screening library from which the two active compounds (**1a**) and (**2a**) were identified.

2. Supporting Table S1: Data collection and refinement summary

3. Supporting Experimental Procedures

- 3.1 Chemical synthesis
- 3.2 Chemical substances
- 3.3 Chemical libraries
- 3.4 Plasmid construction and protein expression
- 3.5 Fluorescence polarization assays and high-throughput screening
- 3.6 Homology modeling and molecular docking
- 3.7 Isothermal Titration Calorimetry
- 3.8 Pin1 Isomerase assay
- 3.9 HSQC-NMR
- 3.10 X-ray crystallography

4. Supporting References

1. Supporting figures

Figure S1

Figure S1: Activity of fosfosal (**1a**) is independent of the presence or absence of DTT and the time of incubation with STAT5b. A) STAT5b was pre-incubated with fosfosal for 1 h in the presence or absence of 1 mM DTT. Subsequently, fluorescein-labeled STAT5b binding peptide 5-carboxyfluorescein-GY(PO_3H_2)LVLDKW was added, and fluorescence polarization was measured either immediately ($t = 0$ h) or after an additional 1 h ($t = 1$ h). B) STAT5b was pre-incubated with fosfosal for the indicated periods of time in the presence of 1 mM DTT. Subsequently, fluorescein-labeled STAT5b binding peptide 5-carboxyfluorescein-GY(PO_3H_2)LVLDKW was added, and fluorescence polarization was measured after an additional 10 min.

Figure S2

Figure S2: Re-synthesized fosfosal is equally potent as the commercially available material.

Figure S3

Figure S3: The binding mode of fosfosal (1) to the STAT5b SH2 domain as predicted by molecular modeling is similar as the binding mode of the phosphotyrosine side chain of a peptide that binds to the STAT1 SH2 domain as determined by X-ray crystallography. The Figure shows an overlay of the crystal structure analysis of Ac-pTyr-Asp-Lys-Pro-His-NH₂ (orange) bound to the SH2 domain of unphosphorylated STAT1 (green) (PDB: 1YVL),¹ and fosfosal (silver) bound to the SH2 domain of STAT5b (grey).

Figure S4

Figure S4: Activity of dexamethasone-21-phosphate (**2a**) is independent of the presence or absence of DTT and the time of incubation with Pin1. A) Pin1 was pre-incubated with dexamethasone-21-phosphate for 1 h in the presence or absence of 1 mM DTT. Subsequently, fluorescein-labeled Pin1 binding peptide 5-carboxyfluorescein-GWFYpSPRLKK-OH was added, and fluorescence polarization was measured either immediately ($t = 0$ h) or after an additional 1 h ($t = 1$ h). B) Pin1 was pre-incubated with dexamethasone-21-phosphate for the indicated periods of time in the presence of 1 mM DTT. Subsequently, fluorescein-labeled Pin1 binding peptide 5-carboxyfluorescein-GWFYpSPRLKK-OH was added, and fluorescence polarization was measured after an additional 10 min.

Figure S5

Figure S5. A) Chemical structures and activities of dexamethasone (**2c**) and dexamethasone-21-acetate (**2d**). B) Neither dexamethasone (**2c**) nor dexamethasone-21-acetate (**2d**) inhibit binding between the peptide 5-carboxyfluorescein-GWFYpSPRLKK-OH and Pin1 in fluorescence polarization assays.

Figure S6

Figure S6: Betamethasone-21-phosphate (**2b**) displays only weak activity against the Pin1 substrate binding domain. A) Binding of **2b** to full-length Pin1 as analyzed by ITC. B) Effect of **2b** on *cis/trans* isomerization of the Pin1 substrate Suc-AEPF-pNA.

Figure S7

Figure S7: Dexamethasone-21-phosphate (**2a**) binds to the substrate binding domain with significantly higher affinity than to the WW domain. A) Binding of **2a** to the Pin1 substrate binding domain (amino acids 46-163) as analyzed by ITC. B) Binding of **2a** to the Pin1 WW domain (amino acids 1-45) as analyzed by ITC.

Figure S8

Figure S8: ^1H , ^{15}N -HSQC-NMR of full-length Pin1 (300 μM) in the absence (red signals) or presence (blue signals) of dexamethasone-21-phosphate (**2a**) (600 μM). The graphic was generated using Sparky.²

8

Figure S10

Figure S10: Alternative depictions of the X-ray structure of the complex of Pin1 (Arg14Ala mutant) and dexamethasone-21-phosphate **2a** (PDB code 3TC5). A) Pin1 Lys63, Arg68, and Arg69 form hydrogen bonds with the phosphate group of **2a**. B) Close-up of the accommodation of the methyl group attached to C-16 and the exocyclic carbonyl group of **2a** by Pin1. C) Electron density of **2a** is depicted as a SigmaA-weighted $2F_o - F_c$ electron density map contoured at 1.5σ .

Figure S11

Figure S11: Chemical structures and names of the 12 natural products or natural product derivatives bearing O-phosphorylated hydroxyphenyl or hydroxymethyl groups within the 5019-member screening library, from which the two active compounds **1a** and **2a** were identified.

2. Supporting Table S1: X-ray data collection and refinement summary

Data collection

Space group	P 3 ₁ 2 1
Cell constants (Å)	
	a = 68.36
	b = 68.36
	c = 79.37
Resolution range (Å)	59-1.4
Wavelength (Å)	1.0
Observed reflections	297502
Unique reflections	42028
Whole resolution range:	
Completeness (%)	98.3
<i>R</i> merge (%)	6.8
<i>I</i> /σ(<i>I</i>)	18.3
Last resolution shell	
Resolution range (Å)	1.7-1.4
Completeness (%)	96.9
<i>R</i> merge (%)	33.8
<i>I</i> /σ(<i>I</i>)	4.67

Refinement

No. of reflections	32347
Resolution (Å)	19.7-1.4
R-factor (%)	17.5
R _{free} (%)	21.1
Average B (Å ²)	15.4
R.m.s bond length (Å)	0.007
R.m.s. angles (Å)	1.146

Content of asymmetric unit

RMSD of monomers (Å)	NA
No. of complexes	1
No. of protein residues/atoms	143/1135
No. of solvent atoms	266

3. Experimental section

3.1 Chemical synthesis

Synthesis of fosfosal (2-(phosphonooxy)benzoic acid (**1a**))

Scheme S1. Reaction conditions: 1) PCl_5 , 2) H_2O .

Salicylic acid (**1f**) (3.85 g, 28 mmol) and PCl_5 (5.5 g, 26 mmol) were stirred at room temperature for 45 min. Subsequently, the reaction was brought to completion by sonication at 60 °C for 90 min. The ice-cooled reaction mixture was dissolved in 10 mL of acetone and 10 mL of benzene, and 1.4 mL (78 mmol) of distilled water was added drop wise. After stirring at 0 °C for 30 min, 20 mL benzene was added and the reaction mixture was allowed to stir at room temperature for 12 h. The precipitate was filtered off, washed with 20 mL benzene, and dried in high vacuum. Yield: 4.7 g (77%). A portion of the crude product was recrystallized from wet acetone/benzene. Upon cooling, white crystals were obtained. Melting point: 160-163 °C. Lit: 162.5-163 °C,³ 153-154 °C.⁴ ^1H -NMR (400 MHz, D_2O): δ = 7.71 (d, J = 8 Hz, 1H), 7.50 (m, 1H), 7.27 (d, J = 8 Hz, 1H), 7.20 (t, J = 8 Hz, 1H) ppm; ^{13}C -NMR (100 MHz, D_2O): δ = 169.90, 150.04 (d, J_{CP} = 7 Hz), 134.26 (d, J = 2 Hz), 131.35, 124.91 (d, J_{CP} = 2 Hz), 123.55 (d, J = 5 Hz), 122.06 (d, J_{CP} = 3 Hz) ppm; ^{31}P -NMR (162 MHz, D_2O): δ = -3.88 ppm. MS $\text{C}_7\text{H}_6\text{O}_6\text{P}$ calcd: 216.99075 [M-H], found: 216.99010.

Synthesis of 3-(phosphonooxy)benzoic acid (**1b**)

Scheme S2. Reaction conditions: 1) PCl_5 , 2) H_2O .

3-hydroxybenzoic acid (**2**) (2.00 g, 14.5 mmol) and PCl_5 (3.02 g, 14.5 mmol) were reacted as described for the synthesis of **1a** to afford **1b** as a white solid. Yield: 1.50 g (48%). Melting point: 189 - 192 °C. ^1H -NMR (600 MHz, $[\text{D}_6]$ -DMSO): δ = 9.14 (bs), 7.73 (m, 2H), 7.48 (m, 1H), 7.38 (m, 1H); ^{13}C -NMR (150 MHz, $[\text{D}_6]$ -DMSO): δ = 167.15, 152.05 (d, J_{CP} = 6 Hz), 132.71, 130.30, 125.33, 125.20 (d, J_{CP} = 8 Hz), 121.14 (d, J_{CP} = 8 Hz) ppm; ^{31}P -NMR (243 MHz, $[\text{D}_6]$ -DMSO): δ = -5.17 ppm. MS $\text{C}_7\text{H}_6\text{O}_6\text{P}$ calcd: 216.99075 [M-H], found: 216.99029.

Synthesis of 4-(phosphonoxy)benzoic acid (**1c**)

Scheme S3. Reaction conditions: 1) PCl_5 , 2) H_2O .

4-hydroxybenzoic acid (**3**) (2.00 g, 14.5 mmol) and PCl_5 (3.02 g, 14.5 mmol) were reacted as described for the synthesis of **1a** to afford **1c** as a white solid. Yield: 1.13 g (38%). Melting point: 192 - 193 °C. Lit: 193 °C.⁵ $^1\text{H-NMR}$ (600 MHz, D_2O): δ = 7.83 (m, 2H), 7.12 ppm (m, 2H); $^{13}\text{C-NMR}$ (150 MHz, D_2O): δ = 167.37, 155.64 (d, J_{CP} = 12 Hz), 131.60, 126.74, 120.35 (d, J_{CP} = 11 Hz) ppm; $^{31}\text{P-NMR}$ (161 MHz, D_2O): δ = -3.96 ppm. MS $\text{C}_7\text{H}_6\text{O}_6\text{P}$ calcd: 216.99075 [M-H], found: 216.99029.

Synthesis of 2-(phosphonomethyl)benzoic acid (**1d**)

Scheme S4. Reaction conditions: i) AIBN, NBS, CCl_4 , 2 h reflux. ii) $\text{P}(\text{OEt})_3$, 3 h reflux. iii) 1 M HCl, 48 h reflux.

Synthesis of 5: Azobisisobutyronitrile (0.548 g, 3.3 mmol) and N-bromosuccinimide (0.593 g, 3.3 mmol) were added to a solution of methyl 2-methylbenzoate (**4**) (0.500 g, 3.3 mmol) in 20 mL dry CCl_4 , and the mixture was refluxed for 2 h. The mixture was quenched with 10 mL of distilled water and extracted twice with ethyl acetate (2x15 mL). The organic phase was dried over Na_2SO_4 , filtered and concentrated. The crude product was purified by column chromatography (10:1 hexane:ethylacetate; R_f = 0.37) to afford **5** as an oil. Yield: 0.67 g (88 %). $^1\text{H-NMR}$ (400 MHz, CDCl_3): δ = 7.97 (dd, J = 7.6 Hz, J = 1.6 Hz, 1H), 7.47 (m, 2H), 7.37 (m, 1H), 4.96 (s, 2H), 3.94 (s, 3H) ppm; $^{13}\text{C-NMR}$ (100 MHz, CDCl_3): δ = 167.15, 139.37, 132.66, 132.81, 131.42, 129.20, 128.65, 52.42, 31.65 ppm; MS $\text{C}_9\text{H}_9\text{BrO}_2$ calcd: 250.97 [M+Na], found: 250.9. Lit: ⁶

Synthesis of 6: Methyl 2-(bromomethyl)benzoate (**5**) (0.450 g, 1.96 mmol) was taken up with 5 mL of triethylphosphite and refluxed for 3 h. After completion of the reaction, 5 mL of distilled water was added and extracted twice with ethyl acetate (2x15 mL). The combined organic phases were dried over Na_2SO_4 , filtered and concentrated. The crude material was purified by column chromatography (3:1 hexane: ethyl acetate; R_f = 0.24) to afford **6** as a colorless oil. Yield: 0.411 g (73 %). $^1\text{H-NMR}$ (400 MHz, CDCl_3): δ = 7.88 (d, J = 7.6 Hz, 1H),

7.42 (m, 1H), 7.37 (m, 1H), 7.28 (m, 1H), 3.97 (m, 4H), 3.88 (s, 3H), 3.78 (d, $J = 22.8$ Hz, 2H), 1.19 (t, $J = 6.8$ Hz, 6H); ^{13}C -NMR (100 MHz, CDCl_3): $\delta = 168.04$, 133.42 (d, $J_{\text{CP}} = 10$ Hz), 132.30 (d, $J_{\text{CP}} = 6$ Hz), 131.98 (d, $J_{\text{CP}} = 3$ Hz), 130.95 (d, $J_{\text{CP}} = 4$ Hz), 130.29 (d, $J_{\text{CP}} = 6$ Hz), 126.96 (d, $J_{\text{CP}} = 6$ Hz), 62.14 (d, $J_{\text{CP}} = 7$ Hz), 52.13, 31.22 (d, $J_{\text{CP}} = 135$ Hz), 16.37 (d, $J_{\text{CP}} = 6$ Hz) ppm; ^{31}P -NMR (162 MHz, CDCl_3): $\delta = 27.32$ ppm; MS $\text{C}_{13}\text{H}_{19}\text{O}_5\text{P}$ calcd: 309.0868 [M+Na], found: 309.0860. Lit: ⁷

Synthesis of 1d: Methyl 2-((diethoxyphosphoryl)methyl)benzoate (**6**) (0.100 g, 62.1 mmol) was refluxed in 5 mL 1M HCl for 48 h. After removal of the solvent *in vacuo*, the raw product was washed with dichloromethane to afford **1d**. Yield: 0.056 g (75%). ^1H -NMR (300 MHz, D_2O): $\delta = 7.69$ (d, $J = 7.8$ Hz, 1H), 7.40 (t, $J = 7.5$ Hz, 1H), 7.24 (m, 2H), 3.50 (d, $J = 22.2$ Hz, 2H) ppm; ^{13}C -NMR (100 MHz, D_2O): $\delta = 172.18$, 133.40 (d, $J_{\text{CP}} = 10$ Hz), 132.65 (d, $J_{\text{CP}} = 4$ Hz), 132.15 (d, $J_{\text{CP}} = 6$ Hz), 130.78 (d, $J_{\text{CP}} = 3$ Hz), 130.56 (d, $J_{\text{CP}} = 6$ Hz), 127.25 (d, $J_{\text{CP}} = 3$ Hz), 32.49 (d, $J_{\text{CP}} = 129$ Hz) ppm; ^{31}P -NMR (162 MHz, D_2O): $\delta = 25.61$ ppm. MS $\text{C}_8\text{H}_9\text{O}_5\text{P}$ calcd: 215.01148 [M-H], found: 215.01147. Lit: ⁷

Synthesis of 2-(difluoro(phosphono)methyl)benzoic acid (**1e**)

Scheme S5. Reaction conditions: i) SOCl_2 , 1 h reflux. ii) BnOH , Py , $0\text{ }^\circ\text{C} \rightarrow 60\text{ }^\circ\text{C}$, 1 h. iii) $\text{BrF}_2\text{C-P(O)(OEt)}_2$, Zn , CuBr , DMF . iv) 3 M HCl , 16 h, reflux.

Synthesis of 9: 2-Iodobenzoic acid (**7**) (0.5 g, 2 mmol) was refluxed with SOCl_2 (0.177 mL, 2.4 mmol) for 1 h. After addition of another 0.5 mL of SOCl_2 , the solution was cooled to room temperature. Excess SOCl_2 was removed at atmospheric pressure. 2-Iodo benzoic acid **8** chloride was isolated by distillation and used in the following reaction without further characterization. To this end, acid chloride **8** was mixed with 3 mL of pyridine at $0\text{ }^\circ\text{C}$, and 0.313 mL of benzyl alcohol were added. This solution was heated to $60\text{ }^\circ\text{C}$ for 1 h and then poured into 10 mL of water. The aqueous layer was extracted with 2x10 mL of diethyl ether. The organic phase was washed with 2x5 mL of 3M HCl , water, saturated aqueous sodium bicarbonate, and water. The ether layer was then dried over NaSO_4 , concentrated and purified by column chromatography (hexane:ethyl acetate 10:1) to afford benzyl 2-iodobenzoate **9** as colorless oil. Yield: 0.625 g (94 %). ^1H -NMR (300 MHz, CDCl_3): $\delta = 8.03$ (dd, $J = 7.9$, 1.1 Hz, 1H), 7.86 (dd, $J = 7.8$, 1.7 Hz, 1H), 7.53 (dd, $J = 7.9$, 1.7 Hz, 2H), 7.42 (m, 4H), 7.18 (dt, $J = 7.83$, $J = 1.7$, 1H), 5.41 (s, 2H); ^{13}C -NMR (75 MHz, CDCl_3): $\delta = 166.32$, 141.38, 135.51, 134.96, 132.74, 131.09, 128.65, 128.57, 128.46, 127.93, 94.25, 67.43 ppm; MS $\text{C}_{14}\text{H}_{12}\text{IO}_2$ calcd: 361.0 [M+Na], found: 360.9. Lit: ⁸

Synthesis of 10: A solution of diethyl (bromodifluoromethyl) phosphonate (0.55 g, 2 mmol) in DMF was slowly added to a suspension of activated zinc dust (0.136 g, 2 mmol) in 3 mL of

dry DMF. The internal temperature was maintained at ~50 °C during the addition, after which the solution was stirred at room temperature for an additional 3 h. Subsequently, CuBr (0.304 g, 2 mmol) was added, and the mixture was stirred at room temperature for 30 min. This was followed by addition of benzyl 2-iodobenzoate **9** (0.350 g, 1 mmol) in 1 mL of DMF and stirring for 38 h at room temperature. Subsequently, 5 mL of water and 5 mL of diethyl ether were added, and the biphasic mixture was passed through celite and extracted with diethyl ether (2x10 mL). The combined organic phases were washed with 10 mL of a saturated NaCl solution, and dried over Na₂SO₄. The crude material was purified by column chromatography (10:1 to 4:1, hexane:ethyl acetate; R_f = 0.28 in 1:1 hexane:ethyl acetate) to afford benzyl 2-((diethoxyphosphoryl)difluoromethyl)benzoate **10** as light yellow oil. Yield: 0.258 g (65 %). ¹H-NMR (400 MHz, CDCl₃): δ = 7.77 (d, J = 8 Hz, 1H), 7.53 (m, 2H), 7.44 (m, 2H), 7.35 (m, 4H), 5.34 (s, 2H), 4.14 (m, 4H), 1.26 (t, J = 6.8 Hz, 6 H) ppm; ¹³C-NMR (75 MHz, CDCl₃): δ = 168.27, 135.55, 132.26 (d, J_{CP} = 3 Hz), 130.74, 130.21, 128.79, 128.74, 128.62, 128.48, 118.28 (dt, J_{CP} = 161 Hz, J_{CF} = 197 Hz), 67.78, 65.05 (d, J_{CP} = 6 Hz), 16.42 (d, J_{CP} = 5 Hz) ppm; ³¹P-NMR (162 MHz, CDCl₃): δ = 2.06 (t, ²J_{PF} = 111.7 Hz) ppm; ¹⁹F-NMR (282 MHz, CDCl₃): δ = -102.06 (d, ²J_{FP} = 111.0 Hz) ppm; MS C₁₉H₂₁F₂O₅P calcd: 399.1 [M+H], found: 399.0.

Synthesis of 1e: 0.2 g (0.5 mmol) of compound **10** was refluxed with 3 mL 3M HCl for 16 h. After removal of the hydrochlorid acid and the volatile cleavage products *in vacuo* at room temperature, 2-(difluoro(phosphono)methyl)benzoic acid **1e** was obtained as an oil. Yield: 0.124 g (98%). For purification, NaHCO₃ (0.098 g, 1.16 mmol, dissolved in 2 mL H₂O) was added to a solution of **10** (0.110 g, 0.43 mmol) dissolved in 2 mL H₂O. Solvents are removed under reduced pressure at room temperature to afford the trisodium salt of **10**. Yield: 0.138 g (99 %). ¹H-NMR (400 MHz, D₂O): δ = 7.95 (d, J = 8 Hz, 1H), 7.32 (m, 2H), 7.16 (d, J = 7.2 Hz, 1H) ppm; ¹³C-NMR (100 MHz, D₂O): δ = 179.72, 138.93 (m), 130.41 (dt, J_{CP} = 14 Hz, J_{CF} = 24 Hz), 129.93, 128.90 (m), 127.23 (m), 126.27, 122.04 (dt, J_{CP} = 181 Hz, J_{CF} = 260 Hz) ppm; ³¹P-NMR (162 MHz, D₂O): δ = 6.37 (t, J = 88.3 Hz) ppm; ¹⁹F-NMR (376 MHz, D₂O): δ = -98.02 (d, J = 88.4 Hz) ppm. Purity by ¹⁹F-NMR: 94 %. The material contains traces of partially cleaved products which cannot be separated from the product due to their polar nature. Longer reaction times resulted in less pure material. Attempts to purify the compound by reversed-phase HPLC were unsuccessful. HRMS C₈H₄F₂Na₃O₅P calcd: 250.99264 [M - 3 Na⁺ + 2H⁺], found: 250.99263.

3.2 Chemical substances

Fosfosal (2-(phosphonooxy)benzoic acid; CAS# 6064-83-1) was purchased from Alfa Aesar (product no. A14417). Dexamethasone (CAS# 50-02-2) was purchased from SERVA (product no. 18660). Dexamethasone-21-acetate (CAS# 1177-87-3; product no. D1881), dexamethasone-21-phosphate (CAS# 2392-39-4; product no. D1159) and betamethasone-21-phosphate (CAS# 151-73-5; product no. B7652) were purchased from Sigma-Aldrich.

3.3 Chemical libraries

The screening library used in this study was assembled from 5 commercially chemical libraries containing known bioactive compounds: 1. Sigma Aldrich Lopac 1280 International Version (1267 compounds); 2. Prestwick Chemical Library (1120 compounds); 3. Microsource Discoveries Spectrum Collection (2000 compounds); 4. Biomol ICCB known bioactives library (480 compounds); 5. Analyticon Discoveries: cherry-picked natural products (1000 compounds). For determination of the number of chemically distinct compounds within these 5 libraries, the sd-files were combined and filtered for duplicate

entries based on the chemical structures of the molecules. This led to 5019 chemically distinct entries in the database.

3.4 Plasmid construction and protein expression

Cloning and expression protocols for STAT1, STAT3, STAT5b, Src SH2, Lck SH2, Plk1 PBD, Plk2 PBD, Chk2 FHA and Pin1 have previously been described.^{9, 10} Coding sequences of human STAT2 (aa 138-702), STAT4 (aa 136-705), and STAT6 (aa 110-651) were amplified by polymerase chain reaction (PCR) from human placenta cDNA and cloned into a modified pET28a plasmid for expression of N-terminal 10x His tagged proteins (STAT2 and STAT6), or into a modified pQE70 vector carrying a C-terminal 6x His-tag and an N-terminal MBP-tag (STAT4). The WW domain (aa 1 – 45) and the PPlase domain of human Pin1 (aa 46 – 163) were amplified by PCR from the full length construct and cloned into a modified pET28a vector. The Pin1 R68, 69A double mutant and the Pin1 Arg14Ala mutant were created by site-directed mutagenesis (Quick Change Kit, Stratagene). Proteins were expressed from *E. coli* Rosetta (Novagen) as previously described,^{9, 10} purified by affinity chromatography, and dialyzed against appropriate buffers. Dialysis buffer for FP assays contained 50 mM Hepes (pH 7.5), 100 mM NaCl, 10 % glycerol, 1 mM EDTA, 1 mM DTT, and 0.1% Nonidet P-40 substitute. Buffer used for ITC dialysis contained 10 mM Hepes (pH 6.8) and 100 mM NaCl. ¹⁵N-labelled Pin1 full-length protein was expressed as described.¹¹ The His-tag of the Pin1 Arg14Ala mutant was removed by digestion with thrombin during dialysis against 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM MgCl₂, 2.5 mM CaCl₂, 1mM β-mercaptoethanol, and 10% (v/v) glycerol for 24 h at 4 °C. Residual His-tagged protein was removed by binding to Ni-NTA; protein without tag was isolated from the flow-through. The protein was fractionated by gel filtration on a Superdex 200 16/60 column (Pharmacia) equilibrated in 10 mM HEPES (pH 7.5), 100 mM NaCl, and 1 mM β-mercaptoethanol. The Arg14Ala Pin1-containing fractions were collected and concentrated up to 18 mg/ml.

3.5 Fluorescence polarization assays and high-throughput screening

Members of the chemical libraries were assayed for their abilities to displace 5-carboxyfluorescein-labeled peptides from their respective interacting protein domains in fluorescence polarization (FP) assays. Peptide sequences were: STAT1: 5-carboxyfluorescein-GY(PO₃H₂)DKPHVL; STAT2: 5-carboxyfluorescein-GpYVFFPSL-OH; STAT3: 5-carboxyfluorescein-GY(PO₃H₂)LPQTV-NH₂; STAT4: 5-carboxyfluorescein-GpYLPQNID-OH; STAT5b: 5-carboxyfluorescein-GY(PO₃H₂)LVLDKW; STAT6: 5-carboxyfluorescein-GpYVPWQDLI-OH; Lck SH2 and Src SH2: 5-carboxyfluorescein-GY(PO₃H₂)EEIP; Plk1 PBD: 5-carboxyfluorescein-GPMQSpTPLNG-OH; Plk2 PBD: 5-carboxyfluorescein-GPMQTSPTPKNG; Chk2 FHA: 5-carboxyfluorescein-GHFDpTYLIRR; Pin1: 5-carboxyfluorescein-GWFYpSPRLKK-OH. Final concentrations: 5-carboxyfluorescein-labeled peptides: 10 nM; proteins: STAT1: 200 nM; STAT2 325 nM; STAT3: 160 nM; STAT4: 60 nM; STAT5b: 120 nM; STAT6: 160 nM; Src SH2: 50 nM; Lck SH2: 40 nM; Plk1 PBD: 100 nM; Plk2 PBD: 80 nM; Chk2 FHA: 160 nM; Pin1: 2000 nM. Peptide and protein stocks were diluted in 10 mM Tris (pH 8.0), 50 mM NaCl, 1 mM EDTA, 0.1 % Nonidet P-40, and 10 % DMSO. Screening was performed in the 384-well format. Fluorescence polarization was assessed in either an Ultra Evolution or Infinite F500 plate reader (both Tecan). Inhibition was calculated based on curve fits using SigmaPlot (SPSS Science Software). IC₅₀ data were converted to K_i-values using the published equation¹² with the following parameters: STAT5b: K_d = 278 nM; Pin1: K_d = 2730 nM.

3.6 Homology modeling and molecular docking

The homology model of human STAT5b based on the amino acid sequence of human STAT5a was generated using HHpred^{13, 14} and MODELLER.¹⁵ Human STAT5a and human STAT5b are 93 % identical on the amino acid level in the SH2 domain. Docking experiments were performed with AutoDock Vina.¹⁶ Default settings were used with the exception of the exhaustiveness, which was set to 50. Docking experiments were run using the STAT5b homology model described above. The docking poses calculated by AutoDock Vina were visualized using PyMOL.¹⁷

3.7 Isothermal Titration Calorimetry (ITC)

ITC experiments were run on a Microcal iTC200 calorimeter. For Pin1 wild-type / dexamethasone-21-phosphate and Pin1 wild-type / betamethasone-21-phosphate, titrations settings were: 20 injections, 22 °C cell temperature, 10 μ cal/s reference power, 60 s initial delay, 0.147 mM Pin1 cell concentration, 2 mM dexamethasone-21-phosphate (**2a**) and 5 mM betamethasone-21-phosphate (**2b**) syringe concentration, respectively, and a stirring speed of 1,000 rpm. The first injection was 0.1 μ L with an injection time of 0.2 s, while all further injections were 2 μ L (4 s injection time). Gaps between injections were 150 s and data points were recorded every 2 s. ITC titrations of Pin1 R68A, R69A with dexamethasone-21-phosphate were run at a protein concentration of 0.171 mM. For the titration of the substrate binding domain, the protein concentration in the cell and the concentration of **2a** in the syringe were 0.253 mM and 4 mM, respectively. Correspondingly, 0.32 mM Pin1 WW domain were titrated with 5 mM **2a**. Data were analyzed using MicroCal Origin 7.0. From all titrations heats of dilution were subtracted and curves were fitted with a one-site binding model. All experiments were performed in triplicate.

3.8 Pin1 Isomerase assay

The *cis/trans* peptidyl-prolyl isomerase activity of Pin1 was determined as described.^{18, 19} Suc-AEPF-*p*NA (Bachem) was used as peptide substrate.²⁰ The peptide substrate was dissolved at 4 mM in dry trifluoroethanol containing 470 mM LiCl. Isomerization assays were measured in buffer containing 50 mM HEPES pH 7.5, 100 mM NaCl, and 5 mM NaN₃.

Solutions and buffers used were pre-cooled to 4 °C. The *cis/trans* peptidyl-prolyl isomerase (PPIase) activity of Pin1 was determined by mixing of 25 μ L 40x concentrated protein solution (dilution series ranging from 500 to 7.8125 nM) with 850 μ L assay buffer. To this end, 100 μ L 60 mg/mL chymotrypsin in assay buffer were added and the reaction was started by addition of 25 μ L peptide substrate. Increasing absorption was monitored at 390 nm at 4 °C in a Beckman Coulter DU500 photometer.

Inhibition of the *cis/trans* peptidyl-prolyl isomerase activity of Pin1 was analyzed by addition of 5 μ L 200x concentrated inhibitor (1:2 dilution series) to the reaction. Pin1 was incubated with the inhibitors for 10 min before the mixtures were transferred to cuvettes and the reactions were started by adding chymotrypsin and Suc-AEPF-*p*NA. The amount of released *para*-nitroanilide (*p*NA) is directly proportional to the *trans* isomer of the peptide substrate starting from a *cis/trans* mixture. The molecular extinction coefficient of *p*NA at 390 nm is 13,300 cm⁻¹ M⁻¹. With the molar extinction coefficient the measured absorption can be correlated to the amount of released *p*NA. Curves were analyzed using a three parameter fit (single) for an exponential rise to maximum in SigmaPlot (SPSS Science Software). Experiments were performed in triplicate.

3.9 ^1H , ^{15}N -HSQC-NMR

All NMR spectra were acquired on a Bruker DRX 600 MHz spectrometer equipped with a cryoprobe. Typically, NMR samples contained up to 0.3 mM of protein in 50 mM Na_2HPO_4 (pH 6.6), 50 mM Na_2SO_4 , and 1mM β -mercaptoethanol. The samples were measured at 300 K. For the ^1H , ^{15}N -HSQC spectrum,²¹ a total of 4096 complex points in t_2 and 128 t_1 increments were acquired. Water suppression was carried out using the WATERGATE 5 sequence. NMR data were processed using Topspin 1.3. Titration experiments were performed using a series of ^1H , ^{15}N -HSQC of ^{15}N -labeled human Pin1 (*hPin1*). The test compounds were dissolved in water. The protein sample containing 10% (v/v) D_2O was titrated to yield the following stoichiometries of protein to compound: 1:0.5; 1:1; 1:1.5; and 1:2. Quantitative analysis of induced chemical shifts was performed on the basis of spectra obtained at saturating conditions. Analysis of ligand-induced shifts was performed by applying the equation of Pythagoras to weighted chemical shifts:

$$\Delta\delta (^1\text{H}, ^{15}\text{N}) = 1000 * \{ [\Delta\delta(^1\text{H})]^2 + [0.2 * \Delta\delta (^{15}\text{N})]^2 \}^{0.5}$$

Amide resonances were categorized according to their weighted chemical shifts using the following cut-off values: weak shift: $\Delta\delta < 0.025$ ppm; medium shift: $0.025 \leq \Delta\delta < 0.075$ ppm; strong shift: $0.075 \text{ ppm} \leq \Delta\delta < 0.125$ ppm; very strong shift: $\Delta\delta \geq 0.125$ ppm. Amino acids with amide resonances displaying weak, medium, strong, and very strong shifts are depicted in light blue, yellow, orange, and red, respectively, in Figure 3A.

3.10 X-ray crystallography

Crystals of human Pin1 Arg14Ala²² were obtained using the sitting-drop vapor-diffusion method. Protein (18 mg/ml) was mixed 1:2 (v/v) with a mother liquor of 2.1-2.4 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M Hepes (pH 7.5), and 1% PEG 400. Crystals appeared within two days at 4 °C. Crystals of ligand-bound complexes were prepared by adding 0.2 μl of a 10 mM dexamethasone-21-phosphate solution (in water) to crystals in mother liquor for a period of 24 h before harvesting. Prior to data collection, all crystals were transferred to a cryosolvent containing 30% glycerol in the mother liquor. The crystals were frozen in liquid nitrogen. Data sets with resolutions of up to 1.4 Å were collected on the SLS beamline PXII at the Paul Scherrer Institute, Villigen, Switzerland. The collected data were integrated, scaled and merged by XDS and XSCALE.²³ The structures were determined by molecular replacement using Molrep²⁴ from the CCP4.²⁵ The structure of the human Pin1 Arg14Ala mutant²² (pdb entry: 2ZR6) was used as probe. Models were refined by Refmac5,²⁶ and rebuilt by MiFit²⁷ and by a subsequent Refmac5 refinement. The refinement was carried out for the resolution range from 19.7-1.4 Å. Signals outside this range did not carry sufficient information to improve the model. Water molecules were added by Arp/Warp.²⁸ The final models were evaluated using Procheck.²⁹ Residues 38-51 are missing in the structure of the complex due to insufficient electron density.

4. Supporting References

- (1) Mao, X.; Ren, Z.; Parker, G. N.; Sondermann, H.; Pastorello, M. A.; Wang, W.; McMurray, J. S.; Demeler, B.; Darnell, J. E., Jr.; Chen, X. (2005) Structural bases of unphosphorylated STAT1 association and receptor binding. *Mol. Cell* 17, 761-771.
- (2) Goddard, T. D.; Kneller, D. G. *SPARKY 3, University of California, San Francisco*.
- (3) Sobotka, H.; Chanley, J. D. (1957) Ortho-carboxy-substituted aromatic phosphates. *US patent 2810741*.
- (4) Brandenberger, H.; Hanson, R. (1953) Eine spektrophotometrische Bestimmungsmethode für saure und alkalische Phosphatasen. *Helv. Chim. Acta* 36, 900-906.
- (5) Chanley, J. D.; Feageson, E. (1955) The mechanism of the hydrolysis of organic phosphates. III. Aromatic phosphates. *J. Am. Chem. Soc.* 77, 4002-4007.
- (6) Eliel, E. L.; Rivard, D. E. (1952) Photobromination of substituted toluenes as a route to substituted benzyl alcohols and benzaldehydes. *J. Org. Chem.* 17, 1252-1256.
- (7) Kreutzkamp, N.; Cordes, G. (1961) Carbonyl and cyanophosphonic acid esters. IV. Derivatives of benzylphosphonic acid. *Arch. Pharm.* 294, 49-56.
- (8) Bak, C.; Praefcke, K. (1978) Aromatic photosubstitution reactions of halogenated benzoic acid benzylesters. *Z. Naturforsch. B* 33, 420-424.
- (9) Reindl, W.; Yuan, J.; Krämer, A.; Strebhardt, K.; Berg, T. (2008) Inhibition of Polo-like Kinase 1 by Blocking Polo-Box Domain-Dependent Protein-Protein Interactions. *Chem. Biol.* 15, 459-66.
- (10) Schust, J.; Sperl, B.; Hollis, A.; Mayer, T. U.; Berg, T. (2006) Stattic: a small-molecule inhibitor of STAT3 activation and dimerization. *Chem. Biol.* 13, 1235-42.
- (11) Bayer, E.; Goettsch, S.; Mueller, J. W.; Griewel, B.; Guiberman, E.; Mayr, L. M.; Bayer, P. (2003) Structural analysis of the mitotic regulator hPin1 in solution: insights into domain architecture and substrate binding. *J. Biol. Chem.* 278, 26183-93.
- (12) Nikolovska-Coleska, Z.; Wang, R.; Fang, X.; Pan, H.; Tomita, Y.; Li, P.; Roller, P. P.; Krajewski, K.; Saito, N. G.; Stuckey, J. A.; Wang, S. (2004) Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. *Anal. Biochem.* 332, 261-73.

- (13) Soding, J.; Biegert, A.; Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucl. Acids Res.* 33, W244-8.
- (14) Soding, J. (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21, 951-60.
- (15) Sali, A.; Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779-815.
- (16) Trott, O.; Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455-61.
- (17) DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*. DeLano Scientific, Palo Alto, CA, USA.
- (18) Kozany, C.; Marz, A.; Kress, C.; Hausch, F. (2009) Fluorescent probes to characterise FK506-binding proteins. *Chembiochem* 10, 1402-10.
- (19) Lu, P. J.; Zhou, X. Z.; Shen, M.; Lu, K. P. (1999) Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science* 283, 1325-8.
- (20) Behrsin, C. D.; Bailey, M. L.; Bateman, K. S.; Hamilton, K. S.; Wahl, L. M.; Brandl, C. J.; Shilton, B. H.; Litchfield, D. W. (2007) Functionally important residues in the peptidyl-prolyl isomerase Pin1 revealed by unigenic evolution. *J. Mol. Biol.* 365, 1143-62.
- (21) Mori, S.; Abeygunawardana, C.; Johnson, M. O.; van Zijl, P. C. (1995) Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J. Magn Reson. B* 108, 94-8.
- (22) Zhang, Y.; Daum, S.; Wildemann, D.; Zhou, X. Z.; Verdecia, M. A.; Bowman, M. E.; Lucke, C.; Hunter, T.; Lu, K. P.; Fischer, G.; Noel, J. P. (2007) Structural basis for high-affinity peptide inhibition of human Pin1. *ACS Chem. Biol.* 2, 320-8.
- (23) Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* 26, 795-800.
- (24) Vagin, A.; Teplyakov, A. (1997) MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Cryst.* 30, 1022-1025.

- (25) Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 50, 760-3.
- (26) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53, 240-55.
- (27) McRee, D. E. (2004) Differential evolution for protein crystallographic optimizations. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2276-9.
- (28) Perrakis, A.; Sixma, T. K.; Wilson, K. S.; Lamzin, V. S. (1997) wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr. D Biol. Crystallogr.* 53, 448-55.
- (29) Morris, A. L.; MacArthur, M. W.; Hutchinson, E. G.; Thornton, J. M. (1992) Stereochemical quality of protein structure coordinates. *Proteins* 12, 345-64.