Supporting Information

Fragment-Based Discovery of the Pyrazol-4-yl urea (AT9283), a Multi-targeted Kinase Inhibitor with Potent Aurora Kinase Activity

Steven Howard*, Valerio Berdini, John A. Boulstridge, Maria G. Carr, David M. Cross, Jayne Curry, Lindsay A. Devine, Theresa R. Early, Lynsey Fazal, Adrian L. Gill, Michelle Heathcote, Sarita Maman, Julia E. Matthews, Rachel L. McMenamin, Eva F. Navarro, Michael A. O'Brien, Marc O'Reilly, David C. Rees, Matthias Reule, Dominic Tisi, Glyn Williams, Mladen Vinković, and Paul G. Wyatt.

Astex Therapeutics Ltd, 436 Cambridge Science Park, Milton Road, Cambridge, CB4 0QA, UK.

Table of Contents

Synthetic procedures and spectral data	S1
Description of analytical and preparative LC/MS systems	S3
Table of LC/MS purity data for target compounds	S4
Method for determination of logD _{7.4}	S5
Experimental details for biological assays	S5
DMPK study methods	S6
nOe method for compound 16	S8
Small molecule X-ray crystallography of compound 16	S10

4-Fluoro-N-[3-(5-morpholin-4-ylmethyl-1H-benzoimidazol-2-yl)-1H-pyrazol-4-yl]-benzamide (9). A mixture of *p*-fluorobenzoic acid (50 mg, 0.36 mmol), aminopyrazole **23** (117 mg, 0.36 mmol), EDC (68 mg, 0.36 mmol) and HOBt (48 mg, 0.36 mmol) was stirred at ambient temperature in DMF (2.5 mL) for 20 h. The mixture was concentrated *in vacuo* and the residue partitioned between EtOAc (5 mL) and water (2 mL). The organic portion washed with saturated aqueous NaHCO₃, dried (MgSO₄) and concentrated *in vacuo*. The residue was purified by SiO₂ flash column chromatography eluting with 0-10% MeOH in EtOAc to give **9** (40 mg, 26%). ¹H NMR (400 MHz, Me-*d*₃-OD): 8.38 (s, 1H), 8.22-8.16 (m, 2H), 7.79-7.70 (m, 1H), 7.61-7.49 (m, 1H), 7.39 (t, J

= 8.5 Hz, 2H), 7.30 (d, J = 8.2 Hz, 1H), 3.75-3.71 (m, 4H), 3.69 (s, 2H), 2.57-2.52 (m, 4H). LC/MS: R_t = 1.29 min, m/z = 421 $[M+H]^+$.

The following three compounds were prepared by following a procedure analogous to that described for compound **10**.

1-(2-Fluorophenyl)-N-[3-(5-morpholin-4-ylmethyl-1H-benzimidazol-2-yl)-1H-pyrazol-4-yl]-urea (11). Purification by preparative LC/MS gave **11** (8.0 mg, 5%). ^1H NMR (400 MHz, Me- d_3 -OD): 8.07 (s, 1H), 7.92 (t, J = 7.9 Hz, 1H), 7.60-7.46 (m, 2H), 7.20 (d, J = 8.2 Hz, 1H), 7.12-6.99 (m, 3H), 3.66-3.61 (m, 6H), 2.49-2.45 (m, 4H); LC/MS: R_t = 1.20 min, m/z = 436 $[M+H]^+$.

1-(2,6-Difluorophenyl)-N-[3-(5-morpholin-4-ylmethyl-1H-benzimidazol-2-yl)-1H-pyrazol-4-yl]-urea (12). Purification by SiO₂ column chromatography eluting with 2-10% MeOH in CH₂Cl₂ gave **12** (30 mg, 39%). ^1H NMR (400 MHz, Me- d_3 -OD): 7.98 (s, 1H), 7.53-7.39 (m, 2H), 7.32-7.21 (m, 1H), 7.15 (d, J = 8.3 Hz, 1H), 6.99 (t, J = 8.2 Hz, 2H), 3.66-3.58 (m, 4H), 3.55 (s, 2H), 2.48-2.34 (m, 4H); LC/MS: R_t = 1.16 min, m/z = 454 $[M+H]^+$.

1-Cyclohexyl-3-[3-(5-Morpholin-4-ylmethyl-1H-Benzoimidazol-2-yl)-1H-Pyrazol-4-yl]-urea (13). Purification by SiO₂ column chromatography eluting with 0-20% MeOH in EtOAc gave **13** (20 mg, 16%). ^1H NMR (400 MHz, Me- d_3 -OD): 8.05 (s, 1H), 7.73-7.45 (m, 2H), 7.26 (d, J = 8.1 Hz, 1H), 3.73-3.69 (m, 4H), 3.68 (s, 2H), 3.63-3.55 (m, 1H), 2.56-2.51 (m, 4H), 2.02-1.95 (m, 2H), 1.80 (d, J = 13.1 Hz, 2H), 1.66 (d, J = 13.0 Hz, 1H), 1.50-1.36 (m, 2H), 1.34-1.22 (m, 3H); LC/MS: R_t = 1.23 min, m/z = 424 $[M+H]^+$.

1-[3-(5-Morpholin-4-ylmethyl-1H-benzoimidazol-2-yl)-1H-pyrazol-4-yl]-3-pyridin-3-yl-urea (14). A mixture of 3-aminopyridine (31.5 mg, 0.33 mmol), Et₃N (0.195 mL, 1.32 mmol) in CH₂Cl₂ (3 mL) was cooled to 0 °C and then treated with triphosgene (85 mg, 0.28 mmol). The reaction was stirred at ambient temperature for 1 h and then aminopyrazole **23** (100 mg, 0.33 mmol) was added and stirring continued until the reaction was complete. The mixture was treated with 2M NaOH in MeOH for 30 minutes and then concentrated *in vacuo*. The residue was purified by SiO₂ column chromatography eluting with 2-20% MeOH in CH₂Cl₂ and then trituration with CH₂Cl₂ followed by Et₂O to give **14** (20 mg, 15%). ^1H NMR (400 MHz, Me- d_3 -OD): 8.75-8.73 (m, 1H), 8.21 (d, J = 4.8 Hz, 1H), 8.18 (s, 1H), 9.10 (d, J = 8.4 Hz, 1H), 7.79-7.49 (m,

2H), 7.41 (dd, J=8.4 Hz, 4.8 Hz, 1H), 7.29 (d, J = 8.0 Hz, 1H), 3.75-3.70 (m, 4H), 3.69 (s, 2H), 2.56-2.52 (m, 4H); LC/MS: R_t = 1.04 min, m/z = 419 [M+H]⁺.

1-[3-(5-Morpholin-4-ylmethyl-1H-benzoimidazol-2-yl)-1H-pyrazol-4-yl]-3-

(tetrahydro-pyran-4-yl)-urea (15). Was prepared following a procedure analogous to that described for compound **16**. Purification by preparative LC/MS gave **15** (7 mg, 8%). ¹H NMR (400 MHz, d₆-DMSO): 12.97 (br, 1H), 12.79-12.68 (m, 1H), 8.94-8.87 (m, 1H), 8.05 (s, 1H), 7.84-7.60 (m, 1H), 7.43-7.31 (m, 2H), 7.19-7.13 (m, 1H), 3.91-3.85 (m, 2H), 3.75-3.65 (m, 1H), 3.61-3.56 (m, 6H), 3.39 (t, J = 12.0 Hz, 2H), 2.42-2.36 (m, 4H), 1.85-1.78 (m, 2H), 1.45 (q, J=12.0 Hz, 2H); LC/MS: R_t = 1.03 min, m/z = 426 [M+H]⁺.

N-[3-(1H-Benzoimidazol-2-yl)-1H-pyrazol-4-yl]-acetamide (6) and N-[3-(1H-Benzimidazol-2-yl)-1H-pyrazol-4-yl]-benzamide (7). Compound **6** and **7** were prepared by following a synthetic route analogous to that described for **22**, **23** and **8**, but starting from benzene-1,2-diamine instead of **21**.

Compound **6**: ¹H NMR (400 MHz, Me-d₃-OD): 8.26 (s, 1H), 7.67-7.62 (m, 2H), 7.30-7.26 (m, 2H), 2.30 (s, 3H). LC/MS: R_t = 0.82 min, m/z = 242 [M+H]⁺.

Compound **7**: ¹H NMR (400 MHz, DMSO-d₆): 13.41-13.11 (m, 2H), 11.61 (s, 1H), 8.42 (s, 1H), 8.11-8.06 (m, 2H), 7.84-7.75 (m, 1H), 7.72-7.67 (m, 3H), 7.59-7.49 (m, 1H), 7.30-7.24 (m, 2H); LC/MS: R_t = 1.33 min, m/z = 304 [M+H]⁺.

Description of LC/MS conditions

Preparative reverse phase mass directed HPLC purification: Compounds were purified using a Waters Fractionlynx system using a linear gradient of aqueous mobile phase (solvent A) and acetonitrile mobile phase (solvent B) as described below. A suitable 3.6 minute gradient was selected according to the analytical trace in order to achieve good separation (e.g. from 5% to 50% B for early retaining compounds; from 35% to 80% B for middle retaining compounds and so on). The purification was performed using either a Phenomenex Synergy MAX-RP or a Phenomenex Luna C18 (2) (10μ, 100 x 21.2mm) column with a flow rate of 24 mL/min, collecting peaks with the correct molecular ion.

Low pH chromatography. Phenomenex Synergy MAX-RP, 10μ, 100 x 21.2mm (alternatively used Thermo Hypersil-Keystone HyPurity Aquastar, 5μ, 100 x 21.2mm

for more polar compounds). Solvent A (H₂O + 0.1% Formic Acid, pH~1.5), Solvent B (CH₃CN + 0.1% Formic Acid).

High pH chromatography. Phenomenex Luna C18 (2), 10 μ , 100 x 21.2mm (alternatively used Phenomenex Gemini, 5 μ , 100 x 21.2mm). Solvent A (H₂O + 10 mM NH₄HCO₃ + NH₄OH, pH=9.2), Solvent B (CH₃CN)

Analytical LC/MS conditions: LC/MS data was obtained on an Agilent Technologies 1200SL-6140 LC-MS system. LC conditions: Acquity UPLC BEH C18; 1.7 μ ; 2.1x50mm column; 0.9 mL/min; gradient 5-95% eluent B over 1.1 minutes. Eluent A = 10mM aqueous NH₄HCO₃+NH₄OH : CH₃CN (95:5) and eluent B = CH₃CN. The column was maintained at a temperature of 50°C. The MS detection was performed with an Agilent 6140 single quadrupole using both positive and negative electrospray ionisation.

Purity Data for Target Compounds

Compound	HPLC method 1 Retention time / min. Purity [%]	Purity / %
5	<i>m/z</i> (ES+) 185 [M+H], R _t = 0.88 min	98.8
6	<i>m/z</i> (ES+) 242 [M+H], R _t = 0.82 min	> 99.0
7	<i>m/z</i> (ES+) 304 [M+H], R _t = 1.33 min	> 99.0
8	<i>m/z</i> (ES+) 403 [M+H], R _t = 1.26 min	> 99.0
9	<i>m/z</i> (ES+) 421 [M+H], R _t = 1.29 min	> 99.0
10	<i>m/z</i> (ES+) 418 [M+H], R _t = 1.19 min	98.0
11	<i>m/z</i> (ES+) 436 [M+H], R _t = 1.20 min	98.0
12	<i>m/z</i> (ES+) 454 [M+H], R _t = 1.16 min	98.2
13	<i>m/z</i> (ES+) 424 [M+H], R _t = 1.23 min	96.3
14	<i>m/z</i> (ES+) 419 [M+H], R _t = 1.04 min	98.5
15	<i>m/z</i> (ES+) 426 [M+H], R _t = 1.03 min	> 99.0
16	<i>m/z</i> (ES+) 382 [M+H], R _t = 1.08 min	> 99.0

Determination of logD_{7.4}. The pKa value of the molecule was first determined using the pH-metric technique. An aqueous solution of the compound was acid-base titrated over a wide pH-range and the pKa calculated from the shape of the titration curve. LogP values were then determined by measuring the shift in pKa when the titration was repeated in a two-phase water-octanol mixture. With these logP and pKa values, a profile of logD versus pH could be drawn, and logD values at specified pH values could be determined by interpolation (Ref. Slater, B.; McCormack, A.; Avdeef, A.; Comer, J. E. A. pH-Metric logP. 4. Comparison of Partition Coefficients Determined by HPLC and Potentiometric Methods to Literature Values. *J. Pharm.Sci.* **1994**, 83, 1280-1283).

CDK2/CyclinA kinase assay. CDK2/cyclin A activity was determined using a radiometric assay to measure the incorporation of γ -phosphate from $\gamma^{33}\text{P}$ -ATP into histone H1. Assay reactions containing 20 mM MOPS pH7.2, 25 mM β -glycerophosphate, 5 mM EDTA, 15 mM MgCl_2 , 45 μM $\gamma^{33}\text{P}$ -ATP (0.78 Ci/mmol), 0.1 mg/mL BSA, 1 mM sodium orthovanadate, 1 mM DTT, 0.12 $\mu\text{g/mL}$ histone H1 and CdK2/cyclin A (produced in-house at Astex Therapeutics) were set up in the presence of compound and allowed to proceed for 5 h. Reactions were then stopped by adding excess phosphoric acid and phosphorylated histone H1 was then separated from excess ATP on a Millipore MAPH filter plate. After washing, scintillant was added and plates counted on a Packard Topcount. IC_{50} values were determined using a sigmoidal dose response equation from Prism GraphPad Software.

CDK1/CyclinB kinase assay. Activity of CDK1/cyclinB was determined as described above for CDK2/CyclinA except that CDK1/CyclinB (Upstate Discovery) was used and reactions were stopped after 2 h.

Other kinase assays. Activity of GSK3- β was determined using a similar format to that used for CDK2. GSK3- β and the substrate phospho-glycogen synthase peptide 2 were obtained from Upstate Discovery. Assay reactions containing enzyme, 5 μM of the above substrate, 10 mM MOPS, pH7, 0.1 mg/mL BSA, 0.001% Brij-35, 0.5% glycerol,

0.2 mM EDTA, 10 mM MgCl₂, 0.01% β-mercaptoethanol and 15 μM γ³³P-ATP were incubated for 3 h. Reactions were then processed as for the CDK2 assays.

Assays for Jak2, 3, Tyk2, FGFR1, 2, 3, VEGFR1, 2, 3, Ret and c-Abl (T315I) were performed in a homogeneous time resolved fluorescence format. For Jak2, Jak3, FGFR1, FGFR2, FGFR3, VEGFR1, VEGFR2, VEGFR3 and Ret a biotinylated Flt3 peptide (VASSDNEYFYVDF) (Alta Bioscience) was used as a substrate. For c-abl (T315I) biotinylated Abltide peptide (EAIYAAPFAKKK) (Millipore) was used as substrate and for Tyk2 a biotinylated TrkC peptide (DVYSTDYRLFN) (Alta Bioscience) was used. In all cases the relevant enzyme and substrate were incubated with test compound and ATP. Reactions were then quenched with EDTA, 2 nM Eu-conjugated anti phospho tyrosine antibody (Perkin Elmer) and 100 nM Streptavidin XL-665 (Cis Bio) added and plates read by time resolved fluorescence (Ex 335 nm, Em 620 and 665 nm) on a Pherastar plate reader (BMG labtech GmbH, Germany). IC₅₀ values were calculated as above.

All other enzyme assays were performed by Millipore (Dundee, UK) using their KinaseProfiler.

Plasma protein binding determination. Plasma protein binding was measure in at least triplicate by equilibrium dialysis. Equilibrium dialysis entailed incubation of compound in mouse plasma containing lithium heparin against phosphate buffer (0.2M, pH 7.4) at 37°C. Chemical stability was confirmed during the incubation period. The resultant plasma samples were diluted with buffer and, using the same ratio, the buffer samples were diluted with plasma giving an equivalent matrix for analysis (described below).

Plasma protein binding was expressed as percentage bound using the equation;

% bound = $((C_p - C_b)/C_p) \times 100$, where C_p is the concentration of compound in plasma, and C_b the concentration in buffer.

***In vitro* intrinsic clearance determination.** The *in vitro* intrinsic clearance of compound **16** was determined in cryopreserved mouse hepatocytes (purchased from In Vitro Technologies, Maryland, USA) using a substrate depletion method. **16** was incubated with 0.3 million cells/ml and the DMSO concentration during the final incubations volume was 0.12%. The reactions were stopped by flash freezing and the

samples extracted with acetonitrile containing an internal standard prior to analysis using the LC-MS/MS method described below. A standard CYP450 substrate (phenacetin) with a medium-high Cl_{int} was used as positive control for the activity of the preparations. Cell viability was assessed using trypan blue over the assay period. Determination of *in vitro* intrinsic clearance from hepatocytes incubation was calculated according to:-

$$In\ vitro\ Cl_{int} = K_{el}\ (1/min) \times \frac{\text{volume of incubation (mL)}}{1 \times 10^6\ \text{viable cells}} = \mu\text{L/min}/1 \times 10^6\ \text{viable cells}$$

Quantitative analysis in support of plasma protein binding, *in vitro* intrinsic clearance, pharmacokinetic and tumour distribution studies. All plasma/buffer samples, plasma samples and tissue homogenates were prepared by protein precipitation using 3 volumes of acetonitrile containing internal standard. Quantification was by LC-MS/MS using one of the systems given below.

LC system	Agilent 1100 series pump (Waters, Manchester, UK) with PAL CTC autosampler (Presearch, Hitchin, UK)	Agilent 1200 series pump and autosampler (Agilent Technologies, Stockport, UK)
Column	Luna C18(2) 3 μ , 30 x 2 mm column with a C18 pre-column filter (Phenomenex, Macclesfield, UK)	Luna HST column (Phenomenex, Macclesfield, UK)
Mobile Phase	0.1% formic acid in either HPLC grade water or HPLC grade acetonitrile (Sigma/ Aldrich, Pool, UK)	10mM ammonium acetate (aq) and 0.1% formic acid in 10% HPLC grade water/90% HPLC grade acetonitrile (Sigma/ Aldrich, Pool, UK)
Gradient	0 min 90:10 aqueous: organic 1 min 5:95 aqueous: organic 2.5min 5:95 aqueous: organic 2.6min 90:10 aqueous: organic 4.5min 90:10 aqueous: organic	0 min 95:5 aqueous: organic 0.25min 95:5 aqueous: organic 1 min 5:95 aqueous: organic 4min 5:95 aqueous: organic 4.1min 90:10 aqueous: organic 6min 90:10 aqueous: organic
Flow Rate	0.8 mL/min	0.6 mL/min
Mass Spectrometer	Micromass Quattro Ultima Pt triple quadrupole (Waters, Manchester, UK)	Agilent Q-ToF (Agilent Technologies, Stockport, UK)

The programme WinNonLin® v4.1 (Pharsight, CA, USA) was used to evaluate pharmacokinetic parameters.

nOe Experiments for compound 16. 4.7 mg of **16** were dissolved in 550 μ L d_6 -DMSO. NMR data were collected at 500MHz using a Bruker DRX500 with a 5mm H,N,C-triple resonance cryoprobe. All spectra were referenced to d_5 -DMSO (multiplet, 2.50ppm). nOe Spectra were collected using the gradient-nOe pulse sequence of Stonehouse et al. (J. Stonehouse, P. Adell, J. Keeler and A.J. Shaka, *J. Am. Chem. Soc.* **116** (1994), pp. 6037–6038). The nOe mixing time was 0.8 seconds; 256 scans were collected for each NOE spectrum.

The numbering scheme used here is shown in Figure 1.

Figure 1

- (a) Compound **16** showing numbering scheme and proposed conformation of the urea. For the purposes of this report the molecule is divided into four separately numbered sections comprising the cyclopropyl-urea, pyrazole, benzimidazole and morpholine groups.
- (b) Details of the cyclopropyl-urea group. H3b and H3b' are on the same side of the cyclopropyl ring as H2 while H3a and H3a' are on the opposing face.

The protons of the cyclopropyl-urea moiety were assigned as follows. Irradiation of the cyclopropyl-urea *H2* proton [m, 1H, 2.61ppm] gave a strong nOe to a two-proton intensity multiplet at 0.83ppm. This multiplet is assigned to the signals of the chemically equivalent protons H3b and H3b'. By elimination, the two-proton intensity

multiplet at 0.55ppm is assigned to the protons H3a and H3a'. Irradiation at 0.55ppm (H3a/H3a') gives a strong nOe to H3b/H3b' as well as to an amide proton at 7.10ppm. This proton is coupled to H2 (data not shown) and is assigned to the N1H proton of the cyclopropyl-urea. This nOe places the N1H proton on the same side of the cyclopropyl ring as Ha/Ha' and indicates that the N1H-C-C-H2 dihedral angle is close to 180°. In addition to nOes to H3b/H3b', H2 also gives a strong nOe to an amide proton at 9.58ppm. This signal is assigned to the remaining amide proton, NH3. The strength of the nOe indicates that H2 and NH3 are considerably less than 4Å apart. This is only possible if the urea is in a *cis-trans* configuration as shown in Figure 1a. This conformation is also observed in X-ray crystal structures of **16** alone and when complexed to Aurora A, and is atypical of ureas. Thus, in solution, as in the solid state, the Aurora-binding conformation of **16** is pre-formed. The observation of an nOe between H3b/H3b' and the H4 and H7* protons of the benzimidazole ring indicates that this conformation may be stabilised by a hydrophobic interaction between these groups. While the cyclopropyl-urea group exists predominantly in its *cis-trans* conformation in DMSO solution, the signals arising from the benzimidazole group indicate the presence of two, slowly interconverting conformations. Each NMR signal of the benzimidazole group is split into two, each of approximately 0.5-proton intensity. These conformations arise from the presence of two equally-populated protamers, one of which is shown in Figure 1(a). The second protamer is shown in Figure 2.

Figure 2: Second protamer of **16** (**16***).

Chemical Shifts of 16 [*d*6-DMSO, 50°C]

Table 1: Protons labelled with an asterisk arise from the alternative benzimidazole protamer, **16***, shown in Figure 2. All chemical shifts are referenced to d5-DMSO at 2.50ppm.

Compound	¹ H NMR Chemical Shifts and Assignments/ppm (50°C, <i>d</i> 6-DMSO, 500MHz)
Compound 16	Cyclopropyl-urea: 9.58(br, 1H, <i>NH3</i>), 7.10(br, 1H, <i>NH1</i>), 2.61(m, 1H, <i>H2</i>), 0.83(m, 2H, <i>H3b/H3b'</i>), 0.55(m, 2H, <i>H3a/H3a'</i>) Pyrazole: 13.01(br, 1H, <i>NH1</i>), 8.05(s, 1H, <i>H5</i>) Benzimidazole: 12.83(br, 0.5H, <i>NH1</i>), 12.81(br, 0.5H, <i>NH1*</i>), 7.55(d, 0.5H, <i>H7*</i>), 7.54(s, 0.5H, <i>H4</i>), 7.42(s, 0.5H, <i>H4*</i>), 7.41(d, 0.5H, <i>H7</i>), 7.17(d, 0.5H, <i>H6</i>), 7.14(d, 0.5H, <i>H6*</i>)

Small Molecule X-ray Crystallography of Compound 16. Crystallographic data were collected using CuK α radiation ($\lambda = 1.5418 \text{ \AA}$) from a Rigaku rotating anode RU3HR, Osmic blue confocal optics, AFC9 $\frac{1}{4} \chi$ goniometer and a Rigaku Jupiter CCD detector. Data collection was controlled by CrystalClear software and images were processed and scaled by Dtrek. Data were corrected using 4th order Fourier absorption correction to compensate for glue and crystal holder (micromount) absorption. The crystal structures were solved using direct methods implemented in SHELXS-97 and refined by SHELXL-97.

Table 1. Crystallographic data.

Compound	16 (dihydrate)	16 (L-lactate salt)
Formula	C ₁₉ H ₂₃ N ₇ O ₂ *2H ₂ O	C ₁₉ H ₂₃ N ₇ O ₂ *C ₃ H ₆ O ₃
T/K	101	97
Space group (#)	P2 ₁ /n (# 14)	P2 ₁ 2 ₁ 2 ₁ (# 19)
a/Å	7.662(10)	9.941(10)
b/Å	15.184(10)	15.034(10)
c/Å	17.711(10)	16.175(10)
β/°	98.53(2)	90
V/Å ³	2038(3)	2417(3)
Z	4	4
Dx/g cm ⁻³	1.361	1.296
μ/mm-1	0.815	0.782
F ₀₀₀	888	1000
R (I>2σ(I))	0.0866	0.0817
wR2 (I>2σ(I))	0.2283	0.2175

Figure 1. TEP representation of **16** dihydrate structure.

Figure 2. TEP representation of **16** (L-lactate salt) structure.