Supplementary Information for

Inhibition of PAD4 activity is sufficient to disrupt mouse and human NET formation.

Huw D. Lewis^{1*}, John Liddle¹, Jim E. Coote², Stephen J. Atkinson¹, Michael D. Barker¹, Benjamin, D. Bax², Kevin L. Bicker⁵, Ryan P. Bingham², Matthew Campbell¹, Yu Hua Chen², Chun-wa Chung², Peter D. Craggs², Rob P. Davis¹, Dirk Eberhard⁴, Gerard Joberty⁴, Kenneth E. Lind³, Kelly Locke², Claire Maller¹, Kimberly Martinod^{6,7}, Chris Patten¹, Oxana Polyakova², Cecil E. Rise³, Martin Rüdiger², Robert J. Sheppard¹, Daniel J. Slade⁵, Pamela Thomas², Jim Thorpe², Gang Yao³, Gerard Drewes⁴, Denisa D. Wagner^{6,8,9}, Paul R. Thompson⁵, Rab K. Prinjha¹ and David M. Wilson^{1, 10}.

Affiliations:

¹ EpiNova DPU, Immuno-Inflammation Therapy Area, GlaxoSmithKline, Medicines Research Centre, Gunnels Wood Road, Stevenage, Herts., SG1 2NY, UK.

² Molecular Discovery Research. GlaxoSmithKline, Medicines Research Centre, Gunnels Wood Road, Stevenage, Herts., SG1 2NY, UK.

³ ELT Boston, Platform Technology and Science, GlaxoSmithKline, Waltham, MA 02451, USA.

⁴ Cellzome GmbH, a GSK company. Meyerhofstrasse 1, 69117 Heidelberg, Germany.

⁵ Department of Chemistry, TSRI, Scripps Florida, 120 Scripps Way, Jupiter, FL 33458, USA.

⁶ Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA 02115, USA.

⁷ Immunology Graduate Program, Division of Medical Sciences, Harvard Medical School, Boston, MA 02115, USA.

⁸ Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA 02115, USA.

⁹ Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA.

¹⁰ Present address: Oncology iMed, AstraZeneca, Alderley Park, SK10 4TG, UK.

Supplementary Figure 1 Fluorescence Polarisation (FP) binding assay characterisation. The FP assay was used to determine affinity for and displacement of GSK215 from PAD4 at varying calcium concentrations: no calcium (open circles), 0.2 mM (filled circles), 2 mM (open squares) and 10 mM (filled squares). (**a**) Structures of fluorescein-labelled FP ligand GSK215 and its unlabelled analogue GSK027. Fluorescein is represented by the black circle. (**b**) GSK215 bound to PAD4 with apparent K_d values of 70, 75, 220 and 770 nM at 0, 0.2, 2 and 10 mM calcium respectively. These calculated $\mathsf{K}_{_{\mathsf{d}}}$ values were then applied to the displacement experiments using free inhibitors in (**c**) and (**d**). (**c**) Representative experiment with IC₅₀ potencies for the displacement of GSK215 by GSK199 calculated as 141, 211, 807 and 2200 nM respectively. (d) Representative experiment with IC₅₀ potencies for displacement of GSK215 by GSK484 calculated as 40, 66, 217 and 576 nM respectively. Mean potencies across >13 replicates for (**c**) and (**d**) are summarised in Figure 1. GSK215 was employed at a concentration of 10 nM throughout these experiments.

Supplementary Figure 2 Reversibility of binding confirmed by mass spectrometry data. Spectra were run following incubation in the absence of (**a**) and presence of 10 mM calcium (**b**). A PAD4 protein mass was observed at 101144 Da. Cl-amidine showed an increase in protein mass at 101418 Da (+ 274 Da for Cl-amidine). The degree of binding was increased in the presence of 10 mM calcium. GSK199, GSK484 and GSK106 all showed no evidence for additional peaks in the presence and absence of calcium, indicating that all three compounds interact reversibly.

Supplementary Figure 3 GSK199 is a reversible PAD4 inhibitor. Complete inhibition of enzyme activity was observed when an excess of GSK199 (100 µM) was incubated with PAD4. Dialysis of this mixture led to an almost complete recovery of PAD4 activity. There was no change in enzyme activity pre- and post-dialysis in the absence of inhibitor.

Supplementary Figure 4 Comparison of binding sites in crystal structures of GSK199 and GSK147 with PAD4. (**a**) 3.29 Å PAD4/GSK199 and (**b**) 3.1 Å PAD4/GSK147 complexes, showing hydrogen bonding interactions as dotted red lines. PAD4 is shown in cyan except for residues 633-645 (yellow), which form a hydrophobic lid with Phe634 and Val643 packing 'on top' of both ligands in this view. (**c**) Structure of GSK147.

Supplementary Figure 5 Experimental electron density maps at compound binding sites (**a**) Stereo of 2Fo-Fc electron density map (1.1 sigma) for GSK147 complex in the region of the ligand (green carbons). Residues 633-644 of PAD4 are highlighted with yellow carbons. (**b**) Stereo Fo-Fc map (3 sigma) of GSK147 complex. The black sphere at the bottom of the figure is a calcium ion. (**c**) Stereo Fo-Fc electron density map (3 sigma) of GSK199 complex. The view is rotated by 45° compared with (**a**) and (**b**), to highlight how the hydrophobic side-chains of Phe634 and Val643 (yellow carbons) pack against the central part of the inhibitor.

Supplementary Figure 6 Experimental electron density maps for novel β -strands. Final electron density map (1.5 sigma) for the two β -strands (634-637 and 642-647) close to GSK199. The map was calculated with phenix.fem. Note that only part of the inhibitor is visible in this view and that the sidechain of Glu642 is not modelled as there was no clear density for it.

Supplementary Figure 7 Alignment of the C-terminal region of the catalytic domain of human PAD family sequences. The Uniprot database accession numbers are; hPAD1 (Q9ULC6), hPAD2 (Q9Y2J8), hPAD3 (Q9ULW8), hPAD4 (Q9UM07) and hPAD6 (Q6TGC4). Sequence numbering is given for hPAD4 only. The alignment includes the sequences for visible residues in three hPAD4 structures: calcium bound hPAD4 (1WDA), calcium free hPAD4 (1WD8) and the GSK199 complex. The RMSD between $C\alpha$ atoms in the structures is shown in the plot. The consensus sequence shows the most common residue in each column of the alignment. For the structures, the secondary structures are represented by red bars (α -helices), yellow arrows (β -strands) and blue bars (turns). The secondary structure annotation for the irregular helical structure encompassing residues 633-640 is represented in this figure as a mixture of helix and turn. For the two liganded structures, residues near the ligand are highlighted in green. Changes can be seen in the structure of residues 631-645 when different ligands are bound. Purple boxes outline the positions of Phe634 and Val643 in hPAD4 and their equivalent residues in the other PAD isoforms. The alignment was generated using the Molecular Operating Environment (MOE), 2012.10; Chemical Computing Group Inc., Montreal.

Supplementary Figure 8 Selectivity of inhibitors for PAD4 over PADs1-3. Stable HEK293 cell pools expressing FLAG-PAD1, FLAG-PAD2, FLAG-PAD3 or FLAG-PAD4 were engineered. Lysates were prepared and incubated at 37 °C in the presence of 2 mM calcium (except the control in the first lane of each blot) and either DMSO alone, 100 µM of GSK199, GSK484 or GSK106 or 200 µM Cl-amidine. Following membrane transfer, peptidyl-citrulline moieties were chemically modified and detected by an antibody capable of recognising this modification across multiple proteins (upper panels). FLAG-PADs were detected by an anti-FLAG antibody (lower panels). Distinct isozyme-dependent citrullination patterns were evident following calcium addition to the lysates, but not in the absence of calcium or in the presence of the pan-PAD inhibitor Cl-amidine. GSK199 and GSK484, but not GSK106, fully inhibited PAD4-induced citrullination but were ineffective in cells expressing PADs 1-3.

Supplementary Figure 9 Residues 633-640 in PAD4 can adopt two different conformations or can be largely disordered. Structural overlay of calcium-free PAD4 (1WD8 – blue) with (**a**) PAD4/benzoyl-L-arginine amide (BAA, green) (PDB:1WDA). (**b**) PAD4 (cyan)/GSK199 (magenta) complex. Region 633-644 is disordered in the calcium-free form (blue dashed line), adopting an irregular helical conformation (orange) in the BAA complex and two b-strands (yellow) in the GSK199 complex. (**c**-**f**) View of BAA and GSK199 complexes with surface representation of the protein and ligand in stick format. These panels highlight the dramatic change in location and shape of the ligand binding pocket following ordering of residues 633-645 in these distinct conformations.

Supplementary Figure 10 Comparison of calcium and compound binding sites. (**a**) Structural overlay of the PAD4/GSK199 complex (blue) with a PAD4/benzoyl-L-arginine amide (BAA) structure (1WDA – orange) which has the substrate analogue inhibitor (BAA) bound at the active site (C645A mutant). The basic nitrogen of GSK199 (arrowed) binds in a similar position to a basic nitrogen from the substrate analogue. Residues 633-645 adopt a very different conformation following binding of GSK199 (including two b-strands – dark blue) compared with that seen in the calcium stabilised substrate binding conformation (yellow – irregular helix in 1WDA structure). (**b**) Structural overlay of the C-terminal region of PAD4/GSK199 complex (blue with pink Ca²⁺ ions) with 1WDA (orange with black Ca²⁺ ions). The extra calcium ion (highlighted by green circle) in the 1WDA structure orders several loop regions (orange – top of figure) in addition to the large conformational change in the active site region (residues 633-645).

Supplementary Figure 11 GSK199 binds to, and inhibits, endogenous PAD4. (**a**) Expression of PAD4 was induced in NB4 promyelocytic leukaemia cells by treating with 1.25% DMSO for 24 h followed by retinoic acid for 24 h. The arrow indicates PAD4 as revealed by immunoblotting using an anti-PAD4 antibody. (**b**) An analogue of GSK199 (CGB208) was covalently attached via amide coupling to activated Sepharose beads and used to retrieve endogenous PAD4 specifically from lysates of differentiated NB4 cells. Lysate from NB4 cells expressing PAD4 was incubated with CGB208 matrix in the presence of DMSO alone (lane 1) or with inhibitors (lanes 2 and 3). The arrow indicates that retrieval of PAD4 by the CGB208 matrix was inhibited in the presence of 10 μ M GSK484, confirming specificity of binding, whereas 10 μ M GSK106 had no effect. Immunoblotting used an anti-PAD4 antibody. (**c**) Lysate from NB4 cells expressing PAD4 was incubated with CGB208 matrix in the presence of DMSO alone or with increasing concentrations of compound as indicated. PAD4 was detected by immunoblotting using anti-PAD4 antibody in a dot-blot format. Capture of PAD4 on CGB208 beads was inhibited in a concentration-dependent manner by GSK199 and GSK484 but not by GSK106. The calculated IC_{50} potencies were 47 nM, 71 nM and >10 μ M, respectively. Results are expressed as mean \pm SD and are

Supplementary Figure 12 Inhibition of cellular citrullination in human neutrophils. (**a**) and (**b**) Lysates from neutrophils incubated with GSK199, GSK484 or GSK106 and stimulated with either calcium ionophore A23187 or *S. aureus* (at an infection ratio of 10:1) were blotted using an antibody against citrullinated histone H3 (upper panels) or total MAPK (lower panels). Calculated IC₅₀ values were 4.6 μ M for GSK199 and 0.4 μ M for GSK484. In both studies, neutrophil viability was unaffected by the inhibitor concentrations used, and probing for unrelated protein (total MAP Kinase) confirmed consistent loading. (**c**) Lysates were blotted with an antibody against modified citrulline. The stimulated global pattern of bands (less pronounced in neutrophils than in over-expressing HEK cells) was inhibited by both active compounds, with GSK484 showing more marked effects than GSK199, and GSK106 lacking inhibitory activity. (**d**) Quantification of the effects of PAD4 inhibitors on cellular levels of H3Cit as assessed by a novel high-content cellular imaging assay in neutrophils stimulated with A23187. Both GSK199 and GSK484 fully inhibited production of H3Cit in a concentrationdependent manner, with IC₅₀ values of 2 μ M and 0.8 μ M, respectively. Results are expressed as mean \pm SEM and are representative of n=3. (e) Representative images from the H3Cit imaging assay confirmed that the dynamic changes in H3Cit levels occurred with no loss of visualised nuclei (evaluated by Hoechst 33342 co-staining) and were therefore not due to differential cell retention or processing.

Supplementary Figure 13 Additional images of ionomycin-stimulated or control mouse neutrophils. Immunostaining for citrullinated histone H3 (green) with DNA counterstaining by Hoechst 33342 (blue). (a) Unstimulated neutrophils. $(b - d)$ Stimulated neutrophils pre-incubated with vehicle, 10 μ M GSK199 or 10 µM GSK484, respectively. White arrowheads indicate NETs identified as H3Cit⁺ cells with spread nuclear morphology. Representative of $n=4-5$. Scale bar = 20 μ m.

Supplementary Figure 14 DNA-stained images of human neutrophils following stimulation with *S. aureus*. (**a**) White arrowheads indicate exemplars used for classification of neutrophil morphology showing lobulated nuclei, delobulated nuclei and diffused NETs. (**b**) Control unstimulated neutrophils with predominantly lobulated nuclear morphology. (**c**) *S. aureus* at an infection ratio of 5 bacteria per neutrophil provoked many neutrophils into becoming diffused NETs. (**d**) GSK484 prevented many neutrophils from forming diffused NETs following *S. aureus* stimulation. (**e**) Control compound GSK106 did not prevent diffused NET induction by *S. aureus*. All NET images (20 per well, n=3 wells per treatment group per donor) were captured by fluorescent imaging. (**f**) Quantification of DNA-stained lobulated nuclei, delobulated nuclei and diffused NETs from 6 donors, using a custom-designed automated algorithm which assessed sixty fields of view over three separate wells per treatment for each donor. (**g**) Comparison of the different treatments on percentage diffused NETs only. The data are derived from the same 6 donors (n=17-18 replicates) quantified in (**f**) and expressed as mean ± SEM of the percentage of NETs counted for each treatment. $*** = p<0.001$ versus vehicle. Scale bar = 20 μ m.

Supplementary Figure 15 Co-staining of DNA and citrullinated histone H3 protein in *S. aureus*stimulated human neutrophils. (**a**) Bacteria alone induced diffused NETs (represented by blue DNA staining in the first column) and the majority of these NETs also positively stained for H3Cit (yellow in the second column; co-staining is represented by white colouration in the third column). (**b**) Treatment with the PAD4 inhibitor GSK484 prior to stimulation resulted in complete abolition of H3Cit from cells while still generating some DNA-stained diffused NET structures. (**c**) Treatment with GSK106 did not abolish the H3Cit staining, suggesting that the GSK484 effect is PAD4-specific. Merged images show NETs stained for both DNA and H3Cit as being white in colour. Scale bar = 20 μ m.

Supplementary Figure 16 PAD4 inhibitors do not affect cellular viability. (**a**) Isolated mouse neutrophils were incubated with GSK199 or GSK484 for 2 h at 37 °C before viability was assessed. Data are expressed as percentage viability compared with untreated control as mean ± SD and are representative of n=4. (**b**) Isolated human neutrophils were incubated with PAD4 inhibitors for 3.5 h at 37 °C before viability was assessed. Data are expressed in arbitrary luminscence units as mean ± SD and are representative of n=3. (**c**) *S. aureus* cultures at exponential phase were incubated with PAD4 inhibitors (all 10 μ M) for 3 h at 37 °C, both with and without shaking, before optical density was measured. Data are expressed in OD units as mean \pm SD and are representative of n=3.

Supplementary Figure 17 Uncut blots from Supplementary Figure 8. (**a**-**b**) Uncut blot for FLAG-PAD1. (**c**-**d**) Uncut blot for FLAG-PAD2. (**e**-**f**) Uncut blot for FLAG-PAD3. (**g**-**h**) Uncut blot for FLAG-PAD4. (**a**), (**c**), (**e**) and (**g**) are the anti-citrulline blots. (**b**), (**d**), (**f**) and (**h**) are the anti-FLAG blots.

Supplementary Figure 18 Uncut blots from Supplementary Figure 11. (**a**) Uncut blot from Supplementary Figure 11(**a**). (**b**) Uncut blot from Supplementary Figure 11(**b**).

Supplementary Figure 19 Uncut blots from Supplementary Figure 12. (**a**) and (**b**) Uncut blots from Supplementary Figure 12(**a**). (**c**) and (**d**) Uncut blot from Supplementary Figure 12(**b**).

Supplementary Table 1. Summary of biochemical potency data from binding and functional assays for PAD4 inhibitors. The FP binding assay was run at a range of calcium concentrations to assess dependency. Replicate numbers are indicated in parentheses. (ND = not determined). As shown by their physicochemical profiles, these compounds occupy optimal physicochemical space and have good developability properties, including high aqueous solubility and permeability.

Supplementary Table 2. Calculated reduced chi² values for modes of inhibition for **GSK199 and GSK484.** Citrulline production was quantified in the presence and absence of inhibitor titrations as described in the online methods section and Lineweaver-Burk plots were used to distinguish between the three possible modes of inhibition: competitive, non competitive and mixed. Visual inspection of the double-reciprocal plots, along with comparison of the reduced chi² values and statistical significances, clearly confirmed mixed inhibition as the best fit for the data. Support for this model comes from FP competition studies (Supplementary Fig. 1), which showed that potency is dependent on the concentration of calcium.

Supplementary Table 3 Data collection and refinement statistics.

*Each datasets was collected from a single crystal. *Values in parentheses are for highest-resolution shell. ##Numbers in brackets are for ligands excluding red (pyrrolopyridine) part of ligand – see top of table.

Supplementary Table 4. Isozyme selectivity of GSK199 and GSK484 for inhibition of PAD family members with respect to BAEE.

The inhibitory effects of GSK199 and GSK484 were assessed on citrulline production by recombinant PAD enzymes. Selective inhibition of PAD4 over the other PAD family members is clear from the K_i and $K_{i'}$ data, which were obtained from global fits of the data to equations representative of mixed inhibition.

Supplementary Table 5. Selectivity data for GSK199, GSK484 and GSK106.

Compounds were screened against a range of protein types using the specified assays.N is the number of replicates.

Supplementary Table 6. Summary of chemoproteomic analysis.

QSSM: Quantified sequence to spectrum matches. QUP: Quantified unique peptides.

