FULL PAPER

Characterisation of the Carboxypeptidase G2 Catalytic Site and Design of New Inhibitors for Cancer Therapy

Dhadchayini Jeyaharan,[a] Carla Brackstone,[a] James Schouten,[b] Paul Davis,[b] and Ann M. Dixon*[a]

Abstract: The enzyme carboxypeptidase G2 (CPG2) is utilized in antibody directed enzyme pro-drug therapy (ADEPT) to catalyze the formation of an active drug from an inert pro-drug. Free CPG2 in the bloodstream must be inhibited before administration of the pro-drug in order to avoid a systemic reaction in the patient. Although a few small-molecule CPG2 inhibitors have been reported, none have been taken forward thus far. This lack of progress is due in part to a lack of structural understanding of the CPG2 active site as well as the absence of small molecules that can block the active site whilst targeting the complex for clearance. The work described here aimed to address both areas. We report the structural/functional impact of extensive point mutation across the putative CPG2 catalytic site and adjacent regions for the first time, revealing residues outside the catalytic region (K208A, S210A and T357A) crucial to enzyme activity. We also describe novel molecules that inhibit CPG2 whilst maintaining accessibility of galactosylated moieties aimed at targeting the enzyme for clearance. This work acts as a platform for future development of high-affinity CPG2 inhibitors that occupy new chemical space and will advance the safe application of ADEPT in cancer treatment.

Introduction

The enzyme carboxypeptidase G2 (CPG2) is a member of the aminoacylase-1 / metallopeptidase 20 (Acy1/M20) family and is considered to be a very promising therapeutic agent in anticancer treatment, specifically in antibody directed enzyme pro-drug therapy (ADEPT)[1]. The concept of ADEPT was put forward in the mid-1980s as a method of restricting the effects of cytotoxic drugs to the site of tumors [1] in an effort to circumvent the adverse impact on healthy tissues observed in treatments such as chemotherapy. In order to exclusively produce and retain a cytotoxic drug at tumor sites, the ADEPT approach utilizes CPG2 to catalyze the degradation of an inert pro-drug into an active drug molecule at the site of a tumor. This can be conceptualized in three stages, as shown in Figure 1A. In the first stage of ADEPT, CPG2 is administered as an immunoconjugate attached to an antibody directed at a tumor-associated antigen. Binding of this antibody to its antigen localizes CPG2 to the tumor, where it is readily available to enzymatically activate a pro-drug (see Figure 1A, Stage 3) [2]. CPG2 was first described as a folate hydrolyzing enzyme, with a preference for folic acid and its analogues, such as methotrexate (MTX) (Figure 1B), displaying an L-glutamate in the carboxy-terminal position [3]. CPG2 mediates the release of the C-terminal glutamate residue in both folic acid and MTX, and has been employed in ADEPT to activate glutamated nitrogen mustard pro-drugs such as ZD2767P (Figure 1B) [4].

Pre-clinical experiments concluded that, between stages 1 and 3, an intermediate step was required involving the inhibition and clearance of unbound CPG2 immunoconjugate from the circulatory system (Figure 1A, Stage 2) [5]. Failure to clear free CPG2 resulted in systemic toxicity, therefore development of CPG2 inhibitors and clearance agents is required for further development of ADEPT as a cancer treatment. One strategy that has been developed to overcome this is the use of a galactosylated antibody selective for the enzyme portion of the immunoconjugate and designed to enhance clearance of the free immunoconjugate through the liver [6]. Galactosylation was selected in order to facilitate rapid uptake by receptors in the liver, minimizing circulatory dwell time, and was shown to reduce CPG2 concentration in the blood without significantly reducing the concentration of enzyme at tumor sites [7-9]. Khan and coworkers took another approach, exploring the use of small molecule inhibitors of CPG2 to inactivate free enzyme in circulation. These were designed to contain the common structural features found in known substrates, such as an α-carboxylate moiety on the glutamate residue and a benzene ring close to the carbonyl carbon of the amide bond [10, 11]. They also included a thiocarbamate moiety attached to a benzene ring (Figure 1B), and while these inhibitors were potent, they were non-competitive towards MTX and have not been reported in applications of ADEPT to date.

Characterization of the active site structure and critical protein-ligand interactions has significant implications in advancing the design of new tumor pro-drugs and CPG2 inhibitors. The crystal structure of un-liganded CPG2 was solved in 1997 (PDB ID 1CG2, shown in Figure 1C) [12], and demonstrated that the enzyme is organized into a discontinuous catalytic domain (residues 23–213 and 326–415) which contains a dinuclear Zn^{2+}-binding site composed of five residues (see Figure 1D). The residues that comprise the CPG2 Zn-binding site are well conserved with other metallopeptidases of known structure, as shown in the multiple sequence alignment of representative proteins of known structure in Figure S1 [12-17]. The dimerization domain, responsible for stabilizing homodimer formation, is inserted into the center of the catalytic domain (residues 214–325). This architecture is also shared with other members of the Acy1 / M20 family including human aminoacylase-1 [18], but further attempts at co-crystallization of CPG2 with ligands, substrates or inhibitors have not been reported. This has led to very little

[a] Dr. D. Jeyaharan, Miss C. Brackstone, Dr. A. Dixon
Department of Chemistry
University of Warwick
Coventry, CV4 7AL, UK
Email: ann.dixon@warwick.ac.uk
[b] Dr. J. Schoulten, Prof. P. Davis
Mologic Ltd
Bedford Technology Park
Thurleigh, Bedford, MK44 2YP, UK

Supporting information or this article is given via a link at the end of the document.

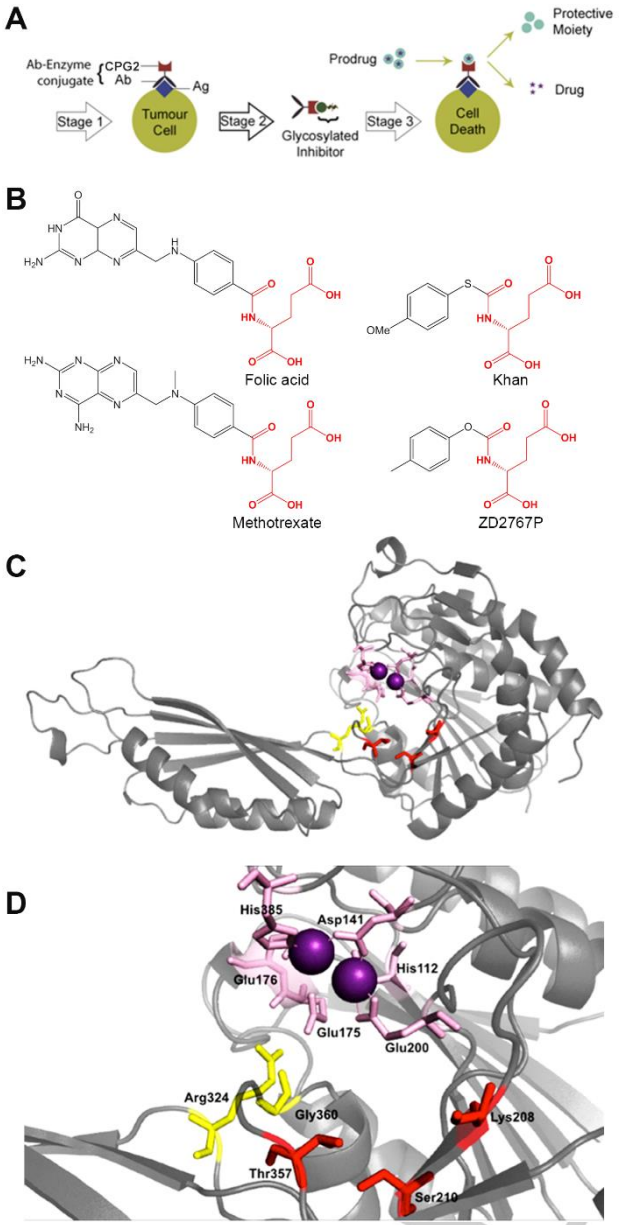


Figure 1. (A) Schematic describing the various stages of antibody-directed enzyme pro-drug therapy (ADEPT). In Stage 1, an antibody-enzyme conjugate (Ab-CPG2) is directed at a tumour-associated antigen and retained at the tumor site. Stage 2 involves administration of a glycoslated enzyme inactivating/clearing agent, which inhibits the enzyme and rapidly removes excess immunoconjugate from the blood via hepatic receptors. Stage 3 describes administration of a non-toxic pro-drug leading to localised generation of a potent cytotoxic agent (e.g. benzoic acid mustards) at tumor sites via CPG2-mediated cleavage of the protective moiety from the pro-drug3. (B) Chemical structure of CPG2 substrates (methotrexate and folic acid) and previously known CPG2 inhibitors (Khan's inhibitor and ZD2767P). Highlighted in red is the D-amino acid glutamate. (C) The different residues within CPG2 proposed to make up the substrate binding-site mapped onto the crystal structure (PDB ID 1CG2). (D) Close-up of the region possibly involved in MTX hydrolysis. In pink are shown residues involved in catalysis, in red residues possibly interacting with the pteridine moiety of MTX (S1 pocket) and in yellow the residues that may be in contact with the glutamate group of MTX (S1' pocket).

progress over the past 20 years in establishing key features of CPG2 such as the location of its substrate-binding site and the presence of additional ligand-binding sites.

Thus far, our current working understanding of the molecular basis of CPG2 activity is largely based on sequence / structural homology to other metallopeptidases and molecular modelling. Specifically, molecular docking and molecular dynamics (MD) simulation was used in 2012 to investigate in silico the possible modes of substrate (specifically MTX) binding to the CPG2 crystal structure [19]. The results from this study suggested that MTX interacts with Glu, His and Asp residues in the putative catalytic domain (shown in pink in Figure 1D). These residues are also involved in Zn^{2+} coordination, and are thus well-conserved throughout the Acy1/M20 family (see Figure S1). Molecular docking also indicated that residues in an adjacent pocket (Pocket S1, shown in red in Figure 1D) could accommodate the N-terminal pteroate moiety of MTX via putative hydrogen bonds between nitrogen atoms in pteridine and the carboxyl oxygens from S210, K208 and T357. The glutamate side chain of MTX was localized in a separate region adjacent to the catalytic site (Pocket S1', shown in yellow in Fig. 1D) and models suggested hydrogen bonding between the glutamate and residues R324 and G360 in the enzyme. Rowsell et al. showed that substitution of R324 to alanine resulted in a mutant CPG2 with very low activity towards MTX [12], confirming that the S1' pocket participates in substrate binding / recognition. Surprisingly, this is the only residue in the proposed MTX binding region that has been studied experimentally via site-directed mutagenesis in vivo or in vitro (although Yachnin and coworkers recently reported creation of CPG2 permutants at sites adjacent to this region [20]). The residues in the putative S1 and S1' pockets appear to be poorly conserved in our sequence alignments of other family members, as shown in Figure S1.

Despite the scarcity of data for CPG2, investigation of homologs including human Acy1 [21] and mouse carnosinase 2 (mCN2) [13] has yielded a model for enzyme function in which substrate binding leads from an "open" conformation to a "closed lid" conformation. This structural rearrangement is required for catalytic activity in several cases, and involves residues in both the catalytic and dimerization domains with key residues in the dimerization domain having been identified [21]. However, we have recently demonstrated that the catalytic domain of CPG2 retains activity against MTX even after deletion of the entire dimerization domain [22]. This result, coupled with the lack of conservation of residues in the S1 and S1' pockets suggests that CPG2 may possess unique substrate-binding characteristics worthy of exploration. To advance our understanding of CPG2 substrate and inhibitor binding, we report for the first time the impact of extensive point mutation across the catalytic site and putative S1 binding pocket on protein folding, thermal stability, and enzyme activity against MTX. This work acts as the first robust test of current, in *silico* models that have represented the state-of-the-art up to this time. We also report development and production of novel molecules that can inhibit the CPG2 catalytic site whilst maintaining the accessibility of galactosylated moieties

aimed at directing the molecules for clearance. Saturation transfer difference NMR spectroscopy (STD-NMR) was used to measure the binding affinities of all three inhibitors to CPG2, as well as identify the binding epitopes within each inhibitor. This work acts as a platform for future development of high-affinity CPG2 inhibitors that occupy new chemical space and will advance the safe application of ADEPT in cancer treatment.

Results and Discussion

Mutagenesis strategy and production of wild-type CPG2 and targeted mutants

Thus far, molecular docking of MTX to CPG2 has given us our best indication of which residues line the substrate binding site, and adjacent hydrophilic and hydrophobic pockets that may direct substrate recognition and cleavage (see Figure 1D). Of the 11 residues highlighted in this previous work, only one residue (R324) has been mutated to demonstrate a reduction in CPG2 activity towards MTX. We recently published results demonstrating high-yield soluble expression, purification and activity of CPG2 lacking the N-terminal signal peptide ($CPG2_{23-415}$) [22]. To test the role of residues within and adjacent to the catalytic site in CPG2 activity, we prepared eight point mutants of the Hisx6-$CPG2_{23-415}$ enzyme in which H112, H385, E175, E200, D141, K208, S210, and T357 were substituted for Ala using site-directed mutagenesis.

The eight resulting proteins were expressed and purified as described previously [22], and the relative expression levels are shown in Figure 2. All eight proteins displayed slow migration in SDS-PAGE, reflecting a molecular weight (~52 kDa) roughly 14% higher than the theoretical mass (45.5 kDa). This anomalous migration has been suggested in the past for extended proteins, acidic proteins [23], and cytosolic proteins that contain surfactant-binding regions [24]. CPG2 is not acidic nor significantly extended, but may contain regions that, when unfolded on the gel, can interact with surfactant molecules slowing migration. Regardless, identity of all proteins was confirmed in later experiments. The H112A mutant had the lowest expression level, followed by K208A and S210A mutants, while the lowest level of expression and the H385A mutant displayed the highest level of expression. The protein yields (per liter of culture) after purification were determined using the BCA assay, and are summarized in Table 1. All substitutions lead to a decrease in expression level (compared to wild type). Sample purity and identity of all purified proteins were confirmed SDS-PAGE gel analyses (Figure S2).

Sample purity and identity were also confirmed using mass spectrometry, and Table 2 summarizes the theoretical versus observed masses of wild-type $CPG2_{23-415}$ (referred to from this point forward as simply CPG2) and all eight mutant proteins. All proteins were identified at the exact predicted molecular mass ±1 Dalton, except for E175A, which was observed at −1.9 Da from the predicted mass. The Hisx6 tag was then removed from all proteins via proteolytic cleavage at a TEV cleavage site followed by further purification as detailed in the Methods.

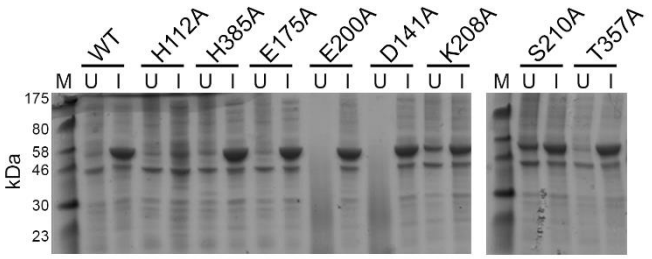


Figure 2. Coomassie-stained SDS-PAGE gel comparing expression levels of wild-type CPG2 and point mutants. Recombinant mutant proteins were tested for expression before (U) and after (I) induction with 0.5 mM IPTG. The dense band migrating just below the 58 kDa marker, which appeared after induction in the majority of cases, was shown to contain the protein of interest. Leaky expression was observed in samples containing the K208A and S210A substitutions, where protein can be seen both pre- and post-induction.

Table 1. Expression yield, enzymatic activity assays and melting temperatures of wild-type CPG2 and eight mutant proteins.

Constructs	Expression yield (mg/L)	Specific Activity (U/mg)	Melting temperature, T_m (°C)
$CPG2_{23-415}$	251.6	96.7	60.4
H112A mutant	23.4	3.9	28.2
H385A mutant	153.1	50.9	55.1
E175A mutant	74.9	2.4	71.4
E200A mutant	125.8	63.7	64.3
D141A mutant	151.0	1.8	53.2
K208A mutant	63.1	3.9	53.4
S210A mutant	114.8	1.5	54.4
T357A mutant	112.1	1.4	55.4

Table 2. Mass average of all expressed constructs, containing N-terminal His_6 tag and TEV cleavage site, obtained from LC-ESI UHR QTOF mass spectrometry.

Constructs	Theoretical Mass (Da)	Observed Mass (Da)	Mass Shift (Da)
$CPG2_{23-415}$	45478.6	45477.9	-0.7
H112A mutant	45412.6	45411.7	-0.9
H385A mutant	45412.6	45411.8	-0.8
E175A mutant	45420.6	45418.7	-1.9
E200A mutant	45420.6	45418.7	-0.9
D141A mutant	45434.6	45433.7	-0.9
K208A mutant	45421.5	45420.8	-0.7
S210A mutant	45462.6	45461.8	-0.8
T357A mutant	45448.6	45447.7	-0.9

Impact of active-site amino acid substitutions on protein fold, thermal stability and activity

The impact of the eight different point mutations on the global protein fold was assessed using circular dichroism (CD) spectrophotometry. As shown in Figure 3A, far-UV CD spectra of the proteins at 25°C have similar shape and intensity, with a negative CD band at 209 nm and positive band at approximately 190 nm. The CD data at 25°C were fitted using the CDSSTR

algorithm in DichroWeb [25], and the percentage secondary structure content was estimated for each protein as shown in Figure 3B. The data indicated that all of the mutated proteins contain very similar secondary structure content to that observed in the wild-type protein.

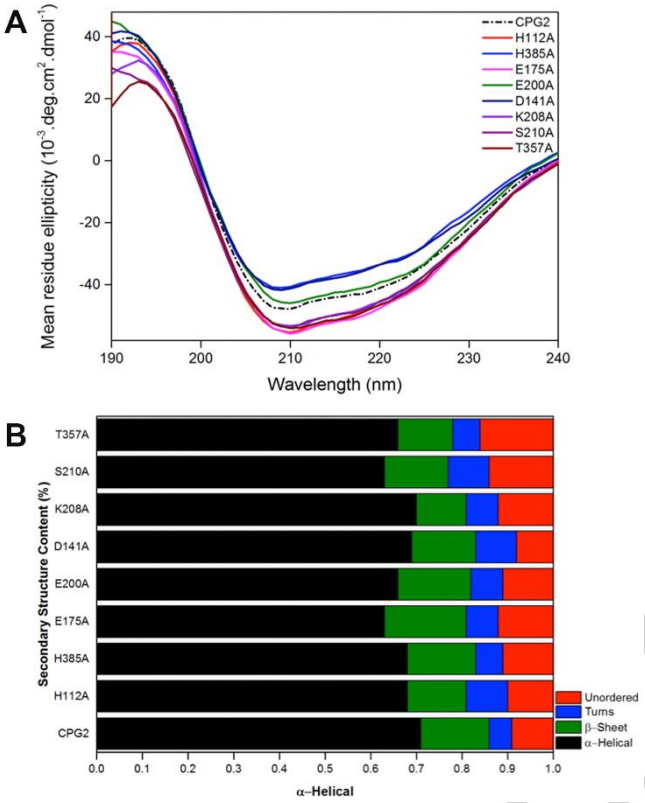


Figure 3. (A) Circular dichroism spectra collected between 190-240 nm for wild type CPG2 and all mutants. (B) Percentage of secondary structure for wild-type CPG2 and all mutants using DichroWeb.[12]

While the global secondary structure was unaffected by the active-site and adjacent substitutions made, the thermal stability of the proteins did show variation (Table 1 and Figure S3). Thermal stability was determined by plotting loss of protein secondary structure (monitored by CD signal at 220 nm) as a function of temperature. All proteins studied here precipitated after they were melted at 90°C, indicating that irreversible unfolding is likely.

The presence of a sigmoidal curve is widely considered a strong indication of the occurrence of a two-state denaturation process [26], where only the folded and unfolded states are reasonably populated at equilibrium [27, 28]. Wild type CPG2 is a non-globular, large multi-domain protein that tends not to exhibit distinct phase transitions during thermal unfolding and, as expected, yielded a sigmoidal relationship suggesting a 2-state transition during thermal denaturation (Figure S3). Fitting of this curve yielded a melting temperature (Tm) of 60.4°C. Proteins containing the H385A, E200A, D141A, K208A, S210A and T357A substitutions also yielded data that fitted well to a sigmoidal curve, and melting temperatures close to that of the wild-type protein (53.2 − 64.3°C). The H112A substitution yielded the lowest Tm value of 28.2°C, suggesting that removal of H112 significantly decreases stability of the protein fold. This result may explain the very low yield of this protein (Figure 2). Conversely, the E175A substitution increased the Tm to 71.4°C, implying that removal of E175 increases stability of the protein.

The catalytic activities of all proteins were estimated by monitoring the rate of degradation of MTX spectrophotometrically at 320 nm (Figure S4A) as described in the Methods. The results, summarized in Table 1, suggest no obvious correlation between thermal stability and activity, therefore differences in activities were interpreted in terms of MTX-binding in the active (and adjacent) site. It was expected that substitutions disturbing the putative catalytic site would abolish activity of CPG2, whereas the mutations involving the adjacent S1 pocket would result in reduced activity of CPG2. This was the case for proteins carrying the H112A, E175A and D141A substitutions (predicted to lie in the catalytic site and coordinate to Zn^{2+}), as all three proteins were rendered inactive against MTX (returning activities of 1.8−3.9 U/mg, values we suggest are insignificant within error and are likely due to non-specific background) likely due to removal of key protein-Zn^{2+} and protein-ligand interactions sites and in line with homologous mutations made in human Acy1 [18]. H385A and E200A (also found in the putative catalytic site) showed specific activities of 50.9 and 63.7 U/mg, respectively, approaching that observed for the wild-type enzyme (96.7 U/mg). Both of these residues are involved with coordination of Zn^{2+} (Figure 1D), so their activity was surprising especially given that the homologous mutations in human Acy1 (H373 and E175) were shown to abolish enzyme activity [18]. This tolerance for the absence of one of these two coordination sites in CPG2 without causing a large loss in catalytic activity or a change in secondary structure, alongside the known requirement for two Zn^{2+} ions across this family of enzymes, suggests that another coordination site is made available upon mutation. Zinc is known to be coordinated by a range of protein side chains, including the nitrogen of histidine, the oxygen of aspartate or glutamate and the sulfur of cysteine, and more rarely to ligands including the hydroxyl of tyrosine [29]. In the case of E200A, inspection of the CPG2 crystal structure [12] reveals that Asp142 is close by in the active site, and upon mutation of E200 to Ala, may be able to rearrange and take up the role of a Zn^{2+} coordination site. In the case of the H385A mutant, the neighboring Tyr384 also in the loop region may coordinate Zn^{2+} to rescue activity. Alternatively, the remaining four residues may coordinate Zn^{2+} in an altered manner, such as the bidentate coordination via both carboxylate oxygens in Glu200 and Glu176.

Substitutions in the adjacent S1 pocket, namely K208A, S210A and T357A, also abolished CPG2 activity (1.4−3.9 U/mg, see Table 1). This almost complete reduction in activity indicates beyond doubt the importance of these adjacent residues in enzyme function, possibly playing a role in substrate recognition. As mentioned earlier, there is little apparent conservation of sequence in the S1 pocket, making broader comparison with other family members difficult. Unno and coworkers reported that

substitution of His228, a residue that resides in the S1 pocket of mouse carnosinase CN2, with Ala also impacted substrate recognition [13]. Okumura and coworkers also reported that H228 was essential to enzymatic reaction [30]. In CN2, the residues in the S1 pocket are quite different to those in CPG2, so it would appear that there is a good deal of sequence diversity in these adjacent regions. Our results, therefore, are the first of their kind in experimentally defining the residues that direct CPG2 catalytic activity, and revealing the unique nature of substrate recognition in this enzyme.

Design, Preparation and Inhibition of Next Generation CPG2 Inhibitors

Antibody directed enzyme pro-drug therapy has been hampered by off-site activation of the pro-drug by circulating protein-enzyme conjugates in normal tissues. The development of a CPG2 inhibitor/clearing agent would facilitate efficient clearance of CPG2 from the bloodstream prior to pro-drug administration, thus eliminating any off-site activation. The ideal agent would contain one component that binds with high affinity to the active site of the CPG2 enzyme, thus blocking activity, and a second component that directs the inactive antibody-enzyme conjugate to the liver for clearance.

Three candidate inhibitors for CPG2 were designed as shown in Figure 4 and synthesized as described in the Methods. These inhibitors all contained the key glutamate residue at the carboxy terminus that directs binding of all known CPG2 substrates and past inhibitors (Figure 1B). Molecular docking results indicated that the Glu side-chain in MTX is stabilized via hydrogen bonding to Arg324 in the S1' pocket, while the amide hydrogen interacts with a carboxyl group on Glu175 in the catalytic site [19]. Our mutagenesis data shows abolishment of activity upon mutation of E175, supporting critical interactions between the Glu moiety and this region of the enzyme. However, L-glutamate was replaced with D-glutamate in order to prevent hydrolysis of the peptide bond whilst retaining binding affinity for CPG2.

Our mutagenesis results also clearly implicated residues in the S1 pocket as key sites of interaction with the substrate. Specifically, substitution of K208, S210 and T357 all abolished activity against MTX. This is in good agreement with the molecular model proposed by Turra and coworkers in which putative hydrogen bonding interactions between the endo and exocyclic nitrogen atoms in the pteridine moiety of MTX and the carbonyl oxygens of Lys 208 and Ser 210 stabilize binding of this region of the molecule. To account for the role of interactions in the S1 pocket, which our mutagenesis studies confirmed played a role in substrate recognition, the molecule termed CP06 contained a naphthylamine group which contains an aromatic heterocyclic component with a nitrogen substitution. This substituent was linked to the glutamate moiety via a peptide bond. A second inhibitor, CP67 (Figure S5), was prepared by addition of three galactose units to the N-terminal naphthylamine group of CP06 via a tripeptide linker. This galactosylated tail was added to facilitate clearance of the CPG2-inhibitor complex via the liver by binding to hepatocyte receptors (i.e. asialoglycoprotein receptors). The third inhibitor tested, referred to as CP42, contained the carboxy terminal glutamate residue, a hydrophobic moiety containing a single aromatic ring (analogous to the Khan inhibitors and nitrogen mustard prodrugs shown in Fig. 1B), and a nitrogen-substituted peptide chain terminating in a single galactose unit (Figure 4A). This inhibitor was designed to test the requirement for a heterocyclic ring versus alternative hydrophobic, nitrogen bearing structures.

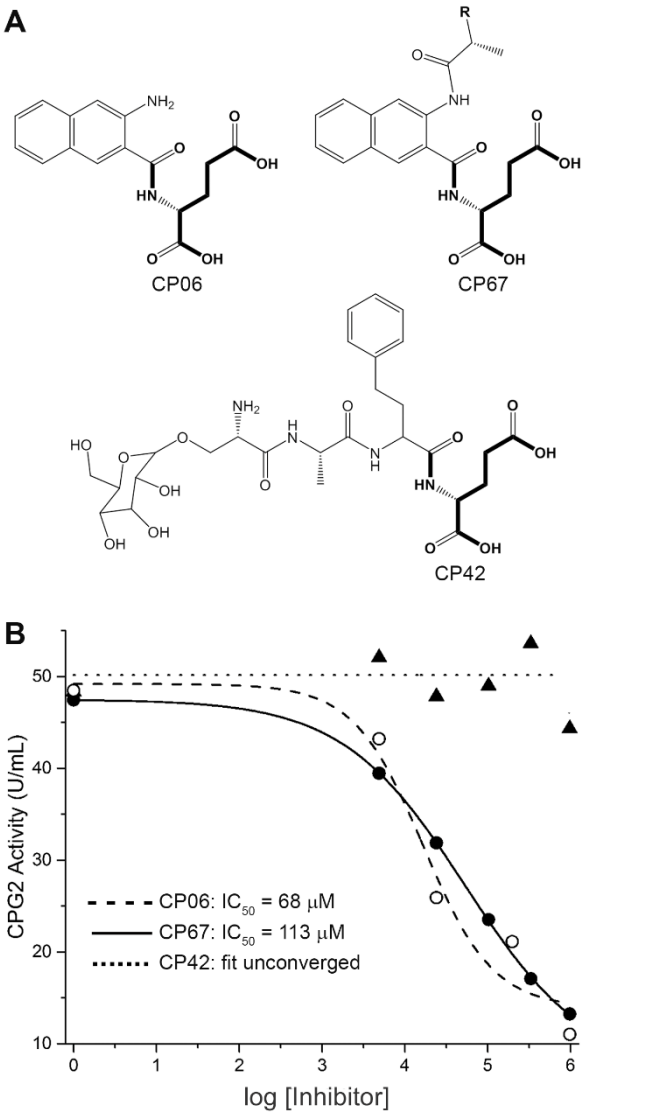


Figure 4. (A) Chemical structures of CP06, CP67 and CP42, with D-glutamate shown in bold. The complete CP67 structure can be found in Figure S4. (B) Plot of CPG2 activity, as determined using the methotrexate cleavage assay (see Methods), in the presence of increasing concentrations of CP06, CP67 and CP42. Data were fitted to an inhibitory dose-response curve in Origin Pro 2015, and the half maximal inhibitory concentration (IC_{50}) was obtained from the fit. IC_{50} values of 68 μM and 113 μM were obtained for CP06 and CP67, respectively, while CP42 showed no inhibition.

The ability of each of these novel inhibitors to inhibit activity of CPG2 was assessed by monitoring the rate of degradation of MTX spectrophotometrically at 320 nm. In order to calculate IC50 values for each inhibitor, the specific activities were measured across a range of inhibitor concentrations as shown in Figure 4B, and fitted to a standard dose-response function using OriginPro. CP06 inhibited the MTX-degrading activity of CPG2 with an IC_{50} of 68 μM, making it the strongest inhibitor of the three prepared here. To demonstrate that this molecule was acting as a genuine inhibitor as opposed to a slowly hydrolysable substrate, we developed a coupled enzyme assay for the detection of glutamic acid. The results are shown in Figure S6 and compared to equivalent data for the well-known CPG2 substrate folic acid. Evolution of glutamate from folic acid is immediately apparent as an increase in fluorescence intensity, and continues to increase steadily over the duration of the measurement. Conversely, the flat line obtained for CP06 suggests that no glutamic acid is evolved in the presence of CPG2 and that it is not a substrate but a genuine inhibitor.

Galactosylation of CP06, to create CP67, increased the IC_{50} by a factor of two (113 μM), suggesting that addition of the tripeptide chain and galactosylated tail weakened binding slightly. Replacing the hydrophobic naphthylamine group with a larger and more complex, but less hydrophobic, moiety as in CP42 eliminated inhibition of CPG2 and the enzyme was able to turn over MTX at its optimal rate.

STD-NMR Measurement of Binding Epitope and Affinity of New Inhibitors

Saturation-transfer difference NMR (STD-NMR) was then used to determine which groups in CP06 and CP67 make up the site of interaction with CPG2, thus facilitating binding and inhibition. In an STD-NMR experiment, selective saturation of a protein host can be transferred to protons in a ligand via the nuclear Overhauser effect if those protons are within 5 Å of the protein [31]. This results in NOE enhancement of the signals from "bound" protons within the ligand, which can be detected via subtraction of the on-resonance 1H spectrum, acquired upon selective irradiation of the protein, from the off-resonance 1H spectrum, acquired with no saturation of the protein. Any proton signals that remain in this difference spectrum identify which protons in the ligand form the binding epitope – this method is thus called "epitope mapping" [32].

The CP06, CP67 and CP42 inhibitors yielded well-resolved 1H 1D NMR spectra at 700 MHz, and these spectra were assigned as shown in Figure 5 (CP06 and CP67) and Figure S7 (CP42). Although some spectral resonances partially overlapped (most notably in the polyether linker in CP67 and galactose-containing regions of both CP67 and CP42), the majority of hydrogen atoms were able to be assigned. The saturation transfer difference (STD) spectrum for each inhibitor in the presence of 20 μM CPG2 is shown immediately below each ^{1}H spectrum in Figures 5 and S7, and clearly revealed efficient saturation transfer

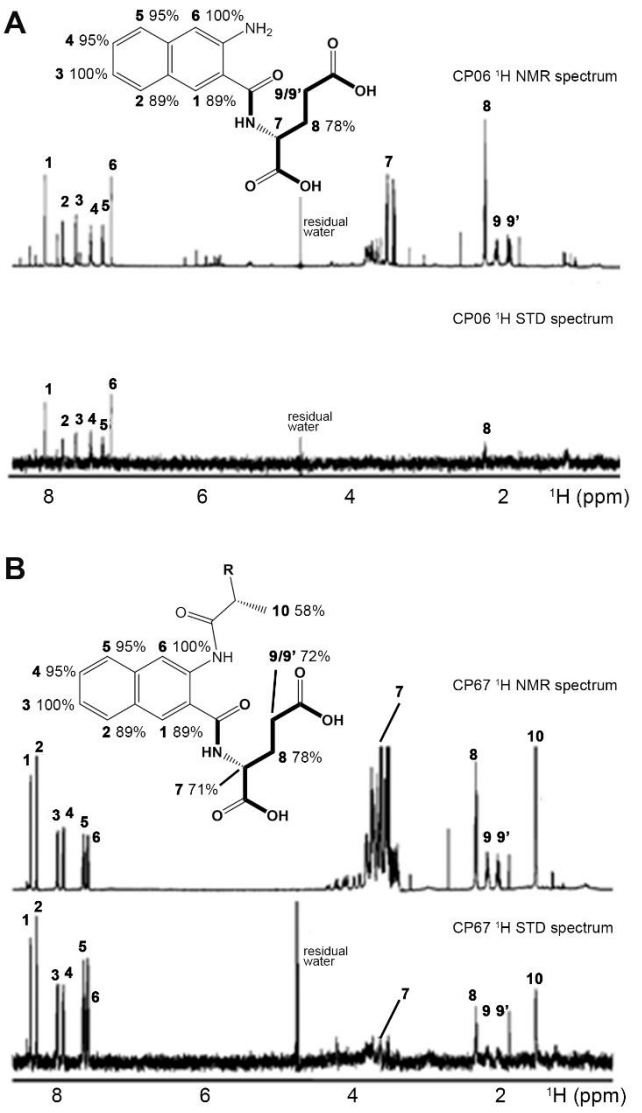


Figure 5. Single Inhibitor titration STD NMR experiments of (A) CP06 and (B) CP67 binding to CPG2. The top spectrum in each panel corresponds to the 1D 1H reference spectrum of the inhibitor, and the bottom to the 1D 1H STD NMR difference spectrum which reveals inhibitor residues that directly interact with CPG2. Also shown are the structures of the inhibitors tested and the relative saturation of the individual protons normalised to that of the H6 proton as determined from 1D STD NMR spectra at an 80 fold molar excess of inhibitor. Absence of the saturation of a proton is not depicted. The few sharp resonances arise from low-molecular-weight are impurities. STD NMR spectra are shown for samples containing 1.6 mM inhibitor and 20 μM CPG2, and were recorded at 298 K on a 700 MHz spectrometer.

from CPG2 to a subset of ^{1}Hs in each inhibitor. The signals contained in the STD spectra identified the protons in each inhibitor that were involved in CPG2 binding – this is one of the unique strengths of STD NMR in comparison with other methods typically used to measure affinity of binding such as isothermal titration calorimetry (ITC) or surface plasmon resonance (SPR). The signal intensities of the individual protons in the STD spectra were normalized to the integral value of the proton displaying the highest STD amplification (which itself was set to 100%), and

values > 35% are shown in Figures 5 and S7. These values were then used to reflect the proximity of each inhibitor proton to CPG2. The lowest STD effects (< 30%) were observed for the galactosylated moieties of CP67 and CP42 as well as the majority of the polyether linker protons, suggesting these regions do not bind to CPG2 and thus remain accessible for hepatocyte receptor binding, as expected. The glutamate protons of CP42 also showed no saturation transfer, which was unsurprising given the complete lack of inhibitory activity observed for this molecule. The data suggest that simultaneous recognition of the carboxy-terminal glutamate and the hydrophobic phenyl moiety within the CPG2 binding site was prevented in this inhibitor, with only the hydrophobic phenyl group acquiring any significant saturation in CP42 likely explained by transient non-specific binding. However, in both CP06 and CP67 high levels of saturation transfer from CPG2 are observed in protons across both the glutamate and the hydrophobic naphthylamine groups (Figures 5A and B), indicating the simultaneous binding of both regions to the enzyme.

STD-NMR experiments can also be used to estimate binding affinity of each inhibitor by obtaining STD spectra for samples with inhibitor:protein molar ratios from 1:1−80:1. The STD binding isotherms, in which STD amplification factor (STD-AF) is plotted against inhibitor concentration, are shown in Figure S8 for all ^{1}Hs displaying saturation transfer. The data were fitted to Equation 2 to yield an estimate of the apparent dissociation constant ($K_{D,app}$) for the interaction, and these values are summarized in Table 3. A range of K_D values were obtained for the various ^{1}Hs in each ligand, leading to a ~30% standard deviation in the average values. This is one of the shortcomings of STD-NMR, and it has been shown that the complex dependence of STD intensities on the spectral properties of the observed signals make it very challenging to obtain accurate affinities directly using this method [33]. In the case of CP06 and CP67 protons, the STD-AF was saturated at high inhibitor concentrations and a good fit to the data was achieved to yield average $K_{D,app}$ values of 330 μM and 240 μM, respectively. However, even at high CP42 concentrations, the STD-AF did not show signs of saturation and suggests that the CP42 inhibitor is only weakly binding to CPG2 (low-mid mM range), if at all. These results are in keeping with the results from the spectrophotometric assay with regards to the lack of CP42 binding and low-mid μM binding of both CP06 and CP67 (Fig 4), however the STD NMR analyses suggest that CP67 binds slightly more strongly than CP06, and also suggest a lower affinity for both inhibitors than was obtained from the IC_{50} values. This is in keeping with a recent report that STD-AF values can often underestimate K_D [33]. For these reasons, STD-NMR data were used to reveal the binding epitopes within each inhibitor, while the absolute values of K_D were interpreted with caution.

Table 3. The $K_{D,app}$ values obtained from each individual curve fit (presented in Figure S6) of STD amplification factor vs. inhibitor concentration for each proton showing saturation transfer. These values were used to calculate an average $K_{D,app}$ (± standard deviation) to estimate a global dissociation constant for each inhibitor.

Inhibitor	Proton number	$K_{D,app}$ (mM) STD-NMR	Average $K_{D,app}$ (mM)
CP06	H1	0.34	0.33 ± 0.10
	H2	0.45	
	H3[a]	1.33	
	H4	0.46	
	H5	0.23	
	H6	0.33	
	H7	0.20	
	H8	high	
	H9	high	
CP67	H1	0.33	0.24 ± 0.08
	H2	0.25	
	H3	0.29	
	H4	0.25	
	H5	0.35	
	H6	0.33	
	H7	0.14	
	H8	0.17	
	H9	0.11	
	H9'	0.08	
CP42	H7	24.67	n/a
	H8	high	
	H9	2.28	

[a] This value was treated as an outlier, beyond 3 × s.d. from the average, and was removed from the calculation of the reported average $K_{D,app}$ and standard deviation.
n/a: not applicable

Conclusions

In this work, we aimed to advance understanding of the molecular determinants of CPG2 activity by providing the first experimental data confirming residues, which drive catalytic activity and ligand recognition. We carried out point mutation of residues predicted (by computational models) to lie in either the catalytic site or an adjacent putative binding pocket. We confirmed that residues H112, E175 and D141 lie in the catalytic site and substitution of these residues with Ala results in abolishment of activity. Two additional residues predicted to lie in the catalytic site and coordinate Zn, namely H385 and E200, appear to hold less critical roles in catalysis as their absence is tolerated by CPG2 with respect to both protein folding and activity. We also demonstrated the key role of residues K208, S210 and T357 (which are adjacent to the catalytic site and proposed to form the S1 pocket) in catalysis for the first time. We propose that, similar to the S1' pocket which has been shown to direct recognition of the glutamate portion of CPG2 substrates, the S1 pocket directs recognition of the pteridine moiety in MTX and the naphthylamine moiety in our CPG2 inhibitors CP06 and CP67 (and distinctly absent in CP42). This relationship between CPG2 binding and structure can help to direct the design of future CPG2 inhibitors with improved affinities.

While comparison of the IC_{50} values obtained here for CP06 and CP67 to previously synthesized CPG2 inhibitors (e.g. the lead thiolate-containing inhibitor reported by Khan and co-workers[6]

with a K_i of 0.3 μM) suggests there is scope for improvement of binding and inhibition, it should be emphasized that strong binding was not our sole aim. The aim here was to create a new type of molecule capable of simultaneously inhibiting CPG2 whilst containing complex oligosaccharide moieties that could direct the enzyme for clearance via the kidneys or liver. CP67 fulfills this aim. STD-NMR epitope mapping confirmed that the glutamate and naphthylamine moieties in CP67 form a viable CPG2 binding-site with μM affinity which is entirely independent from, and does not interfere with, an attached galactosylated domain. Likewise, none of the protons of the galactosylated domain in CP67 revealed proximity to CPG2 in STD-NMR experiments, suggesting this region is readily accessible for hepatocyte receptor binding. CP67 also opens up new chemical space in which to develop novel inhibitors as it contains a naphthylamine group, a functional group that has not been present in any previously reported CPG2 substrate or inhibitor.

Experimental Section

Construction of CPG2 Mutants

The CPG2-encoding gene in the absence of the signal peptide (CPG2 $_{23-415}$, residues 23-415) cloned in plasmid pET-151-D-Topo (Invitrogen, UK) previously developed in our laboratory (pET151-CPG223-415), was used as a template for PCR-based site-directed mutagenesis. The approach used was adapted from the QuikChange® Site-Directed Mutagenesis Kit from Invitrogen (Agilent Technologies, UK). The PCR reaction was performed using the high fidelity Phusion DNA (NEB, UK) polymerase to minimise the chances of introducing unwanted mutations in both the gene and the plasmid DNA backbone, and all mutated genes were verified via sequencing (GATC, Germany).

Overexpression of Wild-type $CPG_{223-415}$

CPG223-415 and all mutated proteins were overexpressed with an N-terminal His6 tag in the pET151 vector using BL21 (DE3) E.coli cells (New England Biolabs, UK). Protein expression was carried out by inoculating 1 L of 2 × YT medium (supplemented with 100 μg L^{-1} ampicillin) with a 100 mL overnight starter culture followed by shaking at 180 rpm and 37°C until an $OD_{600\ nm}$ = 2 was reached. Induction of CPG2 expression was initiated with the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM (unless otherwise stated), and cells were grown at 15°C yielding the highest expression levels. Cells were harvested by centrifugation at 4229 × g for 20 min. at 4°C.

Purification of Recombinant Wild Type and Mutant CPG2 Proteins

Cell pellets were resuspended in 10 ml Tris buffer (20 mM Tris, 137 mM NaCl, 1 mg mL^{-1} lysosome; 1mM 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride; pH 7.3) and cells were lysed using a cell disruptor (30 kPSI). Once lysed, the insoluble fraction was pelleted by centrifugation (47,850 × g, 20 mins, 4°C). The Hisx6-tagged protein of interest was purified from the soluble fraction using Ni-NTA chromatography (Novagen). Imidazole was removed via buffer exchange using a PD-10 column. The N-terminal His-tag was removed by incubating CPG2 with TEV protease at 4°C for 48 hours. Cleaved protein was further purified using Ni-NTA resin (Novagen) followed by FPLC purification on a HiLoad 16/600 Superdex 75 GL column (GE Healthcare UK Ltd. Buckinghamshire, UK) pre-equilibrated with 20 mM Tris HCl buffer (pH 8) containing 100 mM NaCl (Sigma, UK). FPLC was performed at 4°C and a flow rate of 0.5 mL min^{-1}, and protein elution was monitored via absorbance at 280 nm.

Sample Identification and Mass Determination by Mass Spectrometry

Expression of the wild-type $CPG_{223-415}$ and mutant proteins was confirmed using liquid chromatography electrospray ultra-high resolution quadrupole time-of-flight mass spectrometry (LC-ESI UHR QTOF). Samples containing 20 μM protein dissolved in 20 μM ammonium bicarbonate were diluted five-fold with 30% acetonitrile (ACN) with 0.1% formic acid (FA), and 40 μL aliquots of these were loaded onto a C4 column with 100% water and 0.1% FA as solvent A and 100% ACN with 0.1% FA as solvent B using a 3-step gradient (75% of solvent A for 16 min, 100% of solvent B for 5 min and 75% of solvent A for 11.5 min) and eluted directly into the MaXis II QTOF mass spectrometer (Bruker Daltonics). The resulting total ion count data (summed intensity across the entire range of masses being detected at every point in the analysis) was extracted and mass averaged in the region containing the proteins of interest, and deconvoluted using the Data Analysis Software (Bruker Daltonik GmbH; Bremen, Germany).

Enzymatic Activity and Inhibition assay for CPG2 mutants

The activity of CPG2 was estimated by measuring its ability to cleave methotrexate (MTX, which absorbs at 320 nm) into two species with altered absorbance properties, as described previously [22]. In the assay, 50 μL of CPG2 solution (20 mM Tris buffer, 100 mM NaCl, pH 7.3) was added to a 1 mL quartz cuvette containing 900 μL of the assay buffer (0.1 M Tris-HCl, 0.2 mM $ZnCl_2$, pH 7.5) and 50 μL 0.6 mM methotrexate solution (0.1 M Tris-HCl, 0.2 mM $ZnCl_2$, pH 7.5) pre-equilibrated to 37°C before addition of CPG2. When a potential inhibitor of CPG2 was tested, it was first diluted with the assay buffer. In cases where the final concentration depended on solubility, the volume of assay buffer was adjusted to yield a final volume of 1 mL. Absorbance at 320 nm (A320¬) was measured using a Perkin Elmer Lambda 35 UV/VIS Spectrometer with Peltier temperature control unit for 1 minute, in time steps of 2 sec, and absorbance versus time (min) was plotted for each sample. The slope was calculated for the linear portion of the plot where the rate of the absorbance at 320 nm declined (typically between 0-1 min). The activity was obtained in units per millilitre (U mL^{-1}), where one unit is the amount of enzyme required to catalyse the hydrolysis of 1 μmol MTX per minute at 37°C. The measured absorbance of the assay buffer (blank) was subtracted from all measurements.

Amplex® Red Assay for Evolution of Glutamic Acid

Evolution of glutamic acid by CPG2 was measured using the Amplex® Red hydrogen peroxide/peroxidase assay kit (Thermo Fisher, UK). Briefly, if a CPG2 substrate (e.g. folic acid) was added to CPG2, glutamic acid will be released. This glutamic acid was reacted with L-glutamate oxidase to produce α-ketoglutarate and H_2O_2. H_2O_2 was detected via a 1:1 reaction with the Amplex® Red reagent (10-acetyl-3,7-dihydroxyphenoxazine), in

combination with horseradish peroxidase (HRP), to produce the red-fluorescent oxidation product, resorufin. Resorufin emission was monitored at 585 nm using a fluorescence spectrophotometer. This approach was used to monitor glutamate evolution at 37°C in a black 96-well plate for samples containing CPG2 (0.5 U/ml), folic acid (12.5 µM), and the Amplex® Red assay mixture according the manufacturer's instructions. Folic acid was replaced by the CP06 inhibitor to demonstrate the lack of hydrolysis of the inhibitors.

Circular Dichroism

CD spectra of samples containing 0.1-0.2 mg mL-1 protein dissolved in 20 mM Tris buffer (with 100 mM NaCl, 0.2 mM $ZnCl_2$) were recorded in a 0.1 mm path length cylindrical quartz cuvette (Starna, UK) using a J-815 spectrophotometer (Jasco, UK). Spectra were recorded between 190 and 260 nm, with a bandwidth of 2 mm, a scanning speed of 100 nm min-1 and 8 scans. The CD spectrum of the buffer was recorded as a blank and was subtracted from each protein spectrum. The percentage of secondary structure content was estimated from CD spectra using DICHROWEB (UK) and reference dataset 7 [25, 34].

Thermal Denaturation

To record unfolding profiles of wild-type and mutant CPG2 proteins, samples were pre-warmed at 20°C and then slowly heated to 90°C at a rate of 1°C / min. Unfolding was monitored by CD on a Jasco J-1500 spectrophotometer (Jasco, UK) equipped with a Peltier temperature control unit. Melting profiles at a constant wavelength of 220 nm, Θ220(T), were recorded during sample heating in 1 mm cells, with a bandwidth of 2 mm and a scanning speed of 100 nm min^{-1}. Heat-induced changes in solution turbidity indicating protein aggregation were also recorded by the dynode voltage V(T) that was recorded simultaneously with Θ220(T).

CP06, CP67 and CP42 Inhibitor Production

Peptides were produced using a Liberty Microwave Peptide Synthesiser (CEM Microwave Technology Ltd., UK) and a preloaded Fmoc-Glu(OtBu)-TGA resin (Merck Chemicals). To prepare CP67 and CP42, custom made Fmoc-Ala-ANA-OH building blocks were used to avoid acylation of the naphthyl amino group. Reaction cycles were carried out on a 0.1 mM scale in DMF solvent using a 4-fold excess (with respect to resin loading) of 0.2 M Fmoc and side chain-protected amino acid (where appropriate), 0.5 M HBTU and an 8-fold excess of 2 M DIPEA/NMP. The reaction time was 5 minutes at a temperature of 75°C with double couplings used at each step. Deprotections were carried out in 20% Piperidine / DMF at a temperature of 80°C and reaction time of 2 minutes. Once complete, the resin was rinsed on a manifold with dichloromethane and then ethanol, and dried overnight in a vacuum desiccator. Cleavage and final deprotection of the peptide resin was carried out in 20 ml of 95% TFA / 2.5% triisopropylsilane / 2.5% water for 90 minutes. The reaction liquors were dried by rotary evaporation and peptide precipitated in cold diethyl ether. After air drying, the residue was dissolved in 50 / 50 acetonitrile / water and freeze dried for 15 hours to yield an off white solid. To deacetylate the galactose sugars in CP67 and CP42, the peptide was dissolved in dry methanol and the pH adjusted dropwise to pH 10 with 1 M sodium methoxide in dry methanol. After 30 minutes, acetic acid was added to neutralise the reaction and the solvent was removed under a stream of nitrogen. In all cases, crude peptides were dissolved in 5% acetonitrile / water containing 0.1 % TFA and purified by HPLC on an Onyx C18 4.6×100 mm column (Phenomenex) using a buffer A (Water, 0.1% TFA) and buffer B (Acetonitrile 0.1% TFA) gradient of 5% − 100% B over 10 minutes. Pooled fractions were freeze dried to yield a white fluffy solid. Purified peptide was analysed by direct infusion on a ZMD electrospray mass spectrometer (Waters) to confirm expected mass.

Inhibitor-Binding Studies by STD NMR

Wild-type CPG2 was dissolved in 20 mM Tris-d^{11} (pH 7.2) containing 0.1 M NaCl and 0.2 mM $ZnCl_2$ to a final protein stock concentration of 40 µM. All inhibitors were prepared as concentrated (5 mM) stock solutions. All NMR spectroscopy experiments were performed in 3 mm NMR tubes (Bruker, Germany) on a Bruker Avance 700 MHz instrument equipped with a 5 mm cryoprobe. NMR samples were prepared in 180 µL of 100% D_2O buffer. A pseudo-2D version of the STD NMR pulse sequence was used for interleaved acquisition of off- and on-resonance spectra. The on-resonance frequency was set to either -301 Hz (-0.43 ppm) or -700.24 Hz (-1 ppm) depending on the inhibitor, and the off-resonance frequency was set to 30,000 Hz (42.8 ppm). Spectra were referenced to residual water. In all cases, control experiments containing no protein were performed to confirm the lack of direct saturation of inhibitor protons and select the appropriate on-resonance frequency. STD-NMR spectra were acquired with a 2 s saturation time, 3 s relaxation delay, 8 scans, and 4 dummy scans. STD-NMR spectra were collected for samples containing 20 µM CPG2 and increasing concentrations of inhibitor ranging from 0 − 1.6 mM. In the STD−NMR spectra, the STD intensity of a given inhibitor proton was considered to be directly proportional to the fraction of bound inhibitor, f_B [33, 35] as shown in Equation 1:

$$\mathbf{f_B = \frac{[PI]}{[I]} = \frac{(I_0 - I_{sat})}{I_0}}$$

where [PI] is the concentration of the protein-inhibitor complex, [I] is the total concentration of the inhibitor, and I_0 and I_{sat} are the inhibitor proton peak intensities in the absence and presence of protein saturation, respectively. The STD amplification factor, STD−AF, is defined by Mayer and Meyer as the product of f_B and the inhibitor excess (EI = [I]/ [P]) [31, 36]. Plotting STD−AF values at increasing values of [I] gives rise to a Michaelis-Menten hyperbolic dose-response as described by Equation 2:

$$\mathbf{STD - AF_{[I]} = \frac{STD - AF_{max}\,[I]}{[I] + K_D}}$$

where STD−AF_{max} is a dimensionless scaling factor representing the maximum STD amplification factor for the monitored signals. Mathematical fitting yields both parameters K_D and STD−AF_{max}. For each inhibitor concentration, STD−$AF_{[I]}$ build-up curves were recorded. To determine K_D values following the standard procedure, the STD−AF values (at a given saturation time) of an inhibitor proton were plotted as a function of the

concentration of inhibitor, and the resulting data were fitted to Equation 2 to yield the apparent K_D, $K_{D,app}$.

Acknowledgements

This work was supported by a Biotechnology and Biological Sciences Research Council Industrial Case Studentship (BB/I015965/1) awarded to AMD and PD. The authors wish to thank L. Bowsher for helpful discussions on mutagenesis, P.R. Aston (University of Warwick, Coventry, UK) for mass spectrometry and N. Chmel (University of Warwick, Coventry, UK) for circular dichroism assistance.

Keywords: bacterial expression; site-directed mutagenesis; catalytic site; carboxypeptidase G2 inhibitors; saturation transfer difference nuclear magnetic resonance

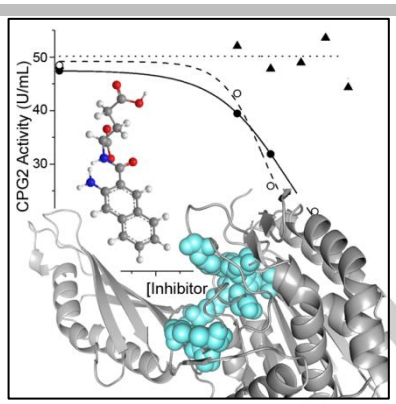
1. K.D. Bagshawe, *Br. J. Cancer*, **1987,** *56*, 531-532.
2. K.D. Bagshawe, F. Searle, C. Springer, J. Boden, R. Melton, R. Sherwood, and M. Jarman, *Br. J. Cancer*, **1989,** *59*, 312-312.
3. C.C. Levy and P. Goldman, *J. Biol. Chem.*, **1967,** *242*, 2933-2938.
4. R.J. Francis, S.K. Sharma, C. Springer, A.J. Green, Hope-Stone, L. D., L. Sena, and R.H.J. Begent, *Br. J. Cancer*, **2002,** *87*, 600–607.
5. K.D. Bagshawe, *Biochem. Soc. Trans.* , **1990,** *18*, 750-752.
6. G.T. Rogers, P.J. Burke, S.K. Sharma, R. Koodie, and J.A. Boden, *Br. J. Cancer*, **1995,** *72*, 1357-1363.
7. S.K. Sharma, K.D. Bagshawe, P.J. Burke, R.W. Boden, and G.T. Rogers, *Br. J. Cancer*, **1990,** *61*, 659-662.
8. S.K. Sharma, K.D. Bagshawe, P.J. Burke, R.W. Boden, G.T. Rogers, C. Springer, R. Melton, and R. Sherwood, *Cancer*, **1994,** *73*, 1114-1120.
9. M.P. Napier, S.K. Sharma, C. Springer, K.D. Bagshawe, A.J. Green, J. Martin, S.M. Stribbling, N. Cushen, D. O'Malley, and R.H.J. Begent, *Clin. Cancer Res.*, **2000,** *6*, 765-772.
10. L.M. DeAngelis, W.P. Tong, S.L. Lin, M. Fleisher, and J.R. Bertino, *J. Clin. Oncol.*, **1996,** *14*, 2145-2149.
11. T.H. Khan, E.A. Eno-Amooquaye, F. Searle, P.J. Browne, H.M. Osborn, and P.J. Burke, *J. Med. Chem.*, **1999,** *42*, 951-956.
12. S. Rowsell, R.A. Pauptit, A.D. Tucker, R.G. Melton, D.M. Blow, and P. Brick, *Structure*, **1997,** *5*, 337-47.
13. H. Unno, T. Yamashita, S. Ujita, N. Okumura, H. Otani, A. Okumura, K. Nagai, and M. Kusunoki, *J. Biol. Chem.*, **2008,** *283*, 27289-27299.
14. B. Chevrier, C. Schalk, H. D'Orchymont, J.M. Rondeau, D. Moras, and C. Tarnus, *Structure*, **1994,** *2*, 283-291.
15. D. Jozic, G. Bourenkow, H. Bartunik, H. Scholze, V. Dive, B. Henrich, R. Huber, W. Bode, and K. Maskos, *Structure*, **2002,** *10*, 1097-1106.
16. J. Badger, J.M. Sauder, J.M. Adams, S. Antonysamy, K. Bain, M.G. Bergseid, S.G. Buchanan, M.D. Buchanan, Y. Batiyenko, J.A. Christopher, S. Emtage, A. Eroshkina, I. Feil, E.B. Furlong, K.S. Gajiwala, X. Gao, D. He, J. Hendle, A. Huber, K. Hoda, P. Kearins, C. Kissinger, B. Laubert, H.A. Lewis, J. Lin, K. Loomis, D. Lorimer, G. Louie, M. Maletic, C.D. Marsh, I. Miller, J. Molinari, H.J. Muller-Dieckmann, J.M. Newman, B.W. Noland, B. Pagarigan, F. Park, T.S. Peat, K.W. Post, S. Radojicic, A. Ramos, R. Romero, M.E. Rutter, W.E. Sanderson, K.D. Schwinn, J. Tresser, J. Winhoven, T.A. Wright, L. Wu, J. Xu, and T.J. Harris, *Proteins*, **2005,** *60*, 787-796.
17. X.-L. Yang, R.J. Skene, D.E. McRee, and P. Schimmel, *Proc. Natl. Acad. Sci. USA*, **2002,** *99*, 15369-15374.
18. H.A. Lindner, V.V. Lunin, A. Alary, R. Hecker, M. Cygler, and R. Menard, *J. Biol. Chem.*, **2003,** *278*, 44496-44504.
19. K.M. Turra, K.F. Pasqualoto, E.I. Ferreira, and D.G. Rando, *J. Mol. Model.*, **2012,** *18*, 1867-1875.
20. B.J. Yachnin and S.G. Khare, *Protein Eng. Des. Sel.*, **2017,** *30*, 321-331.
21. H.A. Lindner, A. Alary, L.I. Boju, T. Sulea, and R. Menard, *Biochemistry*, **2005,** *44*, 15645-15651.
22. D. Jeyaharan, P. Aston, A. Garcia-Perez, J. Schouten, P. Davis, and A.M. Dixon, *Protein Expres. Purif.*, **2016,** *127*, 44-52.
23. V.S. Alves, D.C. Pimenta, E. Sattlegger, and B.A. Castilho, *Biochem. Biophys. Res. Comm.*, **2004,** *314*, 229-234.
24. Y.H. Shi, R.A. Mowery, J. Ashley, M. Hentz, A.J. Ramirez, B. Bilgicer, H. Slunt-Brown, D.R. Borchelt, and B.F. Shaw, *Prot. Sci.*, **2012,** *21*, 1197-1209.
25. L. Whitmore and B.A. Wallace, *Nucleic Acids Res.*, **2004,** *32*, 668-673.
26. P.L. Privalov and N.N. Khechina, *J. Mol. Biol.*, **1974,** *86*, 665-684.
27. C.N. Pace, G.R. Grimsley, S.T. Thomas, and G.I. Makhatadze, *Prot. Sci.*, **1999,** *8*, 1500-1504.
28. J.C. Waldner, S.J. Lahr, M.H. Edgell, and G.J. Pielak, *Biopolymers*, **1999,** *49*, 471-479.
29. K.A. McCall, C. Huang, and C.A. Fierke, *J. Nutr.*, **2000,** *130*, 1437S–1446S.
30. N. Okumura, J. Tamura, and T. Takao, *Prot. Sci.*, **2016,** *25*, 511-522.
31. M. Mayer and B. Meyer, *Angew. Chem. Int. Ed.*, **1999,** *38*, 1784-1788.
32. J.L. Wagstaff, S.L. Taylor, and M.J. Howard, *Mol. Biosyst.*, **2013,** *9*, 571-577.
33. J. Angulo, P.M. Enriquez-Navas, and P.M. Nieto, *Chem. Eur. J.*, **2010,** *16*, 7803-7812.
34. L. Whitmore and B.A. Wallace, *Biopolymers*, **2008,** *89*, 392-400.
35. A. Viegas, J.o. Manso, F.L. Nobrega, and E.J. Cabrita, *J. Chem. Educ.*, **2011,** *88*, 990-994.
36. B. Meyer and T. Peters, *Angew. Chem. Int. Ed.*, **2003,** *42*, 864-90.

Entry for the Table of Contents (Please choose one layout)

FULL PAPER

Text for Table of Contents

The structural and functional impact of extensive point mutation across the entire putative CPG2 catalytic site and adjacent regions (shown in structure) is reported for the first time, revealing that residues outside the catalytic region may direct substrate recognition. NMR was then used to map the binding epitopes and affinities of several novel CPG2 inhibitors that occupy new chemical space and show promise as a platform for future development in cancer treatment.



*Dhadchayini Jeyaharan, Carla Brackstone, James Schouten, Paul Davis, and Ann M. Dixon**

Page No. – Page No.

Characterisation of the Carboxypeptidase G2 Catalytic Site and Design of New Inhibitors for Cancer Therapy