the FEBS Journal

Structural and biochemical characterization of the folyl-poly-γ-L-glutamate hydrolyzing activity of human glutamate carboxypeptidase II

Michal Navrátil[1,2], Jakub Ptáček[2,3], Pavel Šácha[1,2], Jana Starková[1], Jacek Lubkowski[4], Cyril Bařinka[3] and Jan Konvalinka[1,2]

1 Gilead Sciences and IOCB Research Centre, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Prague 6, Czech Republic
2 Department of Biochemistry, Faculty of Sciences, Charles University in Prague, Czech Republic
3 Institute of Biotechnology, Academy of Sciences of the Czech Republic, Prague 6, Czech Republic
4 Center for Cancer Research, National Cancer Institute, Macromolecular Crystallography Laboratory, Frederick, MD, USA

Keywords
arene-binding site; crystal structure; folate hydrolase 1; H475Y(1561C→T) polymorphism; zinc metalloprotease

Correspondence
J. Konvalinka, Gilead Sciences and IOCB Research Centre, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo n. 2, 166 10 Prague, Czech Republic
Fax: +420 220 183 578
Tel: +420 220 183 218
E-mail: konval@uochb.cas.cz
Website: http://www.uochb.cz/web/structure/185.html

(Received 20 February 2014, revised 16 April 2014, accepted 21 May 2014)

doi:10.1111/febs.12857

In addition to its well-characterized role in the central nervous system, human glutamate carboxypeptidase II (GCPII; Uniprot ID Q04609) acts as a folate hydrolase in the small intestine, participating in the absorption of dietary poly-glutamylated folates (folyl-*n*-γ-L-glutamic acid), which are the provitamin form of folic acid (also known as vitamin B_9). Despite the role of GCPII as a folate hydrolase, nothing is known about the processing of polyglutamylated folates by GCPII at the structural or enzymological level. Moreover, many epidemiologic studies on the relationship of the naturally occurring His475Tyr polymorphism to folic acid status suggest that this polymorphism may be associated with several pathologies linked to impaired folate metabolism. In the present study, we report: (a) a series X-ray structures of complexes between a catalytically inactive GCPII mutant (Glu424Ala) and a panel of naturally occurring polyglutamylated folates; (b) the X-ray structure of the His475Tyr variant at a resolution of 1.83 Å; (c) the study of the recently identified arene-binding site of GCPII through mutagenesis (Arg463Leu, Arg511Leu and Trp541Ala), inhibitor binding and enzyme kinetics with polyglutamylated folates as substrates; and (d) a comparison of the thermal stabilities and folate-hydrolyzing activities of GCPII wild-type and His475Tyr variants. As a result, the crystallographic data reveal considerable details about the binding mode of polyglutamylated folates to GCPII, especially the engagement of the arene binding site in recognizing the folic acid moiety. Additionally, the combined structural and kinetic data suggest that GCPII wild-type and His475Tyr variant are functionally identical.

Database
Structural data have been deposited in the RCSB Protein Data Bank database under accession numbers 4MCS (GCPII-His475Tyr-glutamate), 4MCP (GCPII-Glu424Ala-FolGlu_1), 4MCQ (GCPII-Glu424Ala-FolGlu_2) and 4MCR (GCPII-Glu424Ala-FolGlu_3).

Abbreviations
2-PMPA, 2-(phosphonomethyl)-pentanedioic acid; 3D, three-dimensional; ABS, arene-binding site; ARM-P2/4/8, urea based inhibitor containing the arene-binding site targeted dinitrophenyl moiety linked via 2/4/8 ethyleneglycol units; FolGlu_n, folyl-*n*-γ-L-glutamic acid; GCPII, human glutamate carboxypeptidase II; MeO-P4, urea based inhibitor with four ethyleneglycol units containing no arene-binding site targeted moiety; NAAG, N-acetyl-L-aspartyl-L-glutamate; PDB, Protein Data Bank; rhGCPII, recombinant human glutamate carboxypeptidase II.

Introduction

Glutamate carboxypeptidase II (GCPII; EC 3.4.17.21) is a 750 amino acid type II transmembrane glycoprotein [1] and a Zn^{2+}-dependent metalloprotease of the M28 peptidase family (Fig. 1). This enzyme is also known as prostate-specific membrane antigen, folate hydrolase 1, folyl-poly-γ-glutamate carboxypeptidase and N-acetylated-α-linked acidic dipeptidase. These different designations reflect the various functions and tissue distribution of this protein.

The well-known and thoroughly studied enzymatic activity of GCPII is the cleavage of peptide neurotransmitter N-acetyl-L-aspartyl-L-glutamate (NAAG) into N-acetyl-L-aspartate and L-glutamate [2–6]. By contrast, nothing is known about the folyl-poly-γ-L-glutamate hydrolyzing activity of GCPII. At the lumenal surface of the human jejunum, GCPII cleaves γ-linked L-glutamates from folyl-poly-γ-L-glutamic acids ($FolGlu_n$) [7], the storage form of folic acid that cannot pass through the cell membrane [8]. (Structures of the folyl-poly-γ-L-glutamic acids and their building block pteroic acid, which is composed of a substituted pteridine double-ring and p-aminobenzoyl moiety, are shown in Fig. 1) Atlhough the structural and enzymological aspects of $FolGlu_n$ cleavage have not been characterized and folate is crucial for replication of rapidly dividing cells, the physiological relevance and

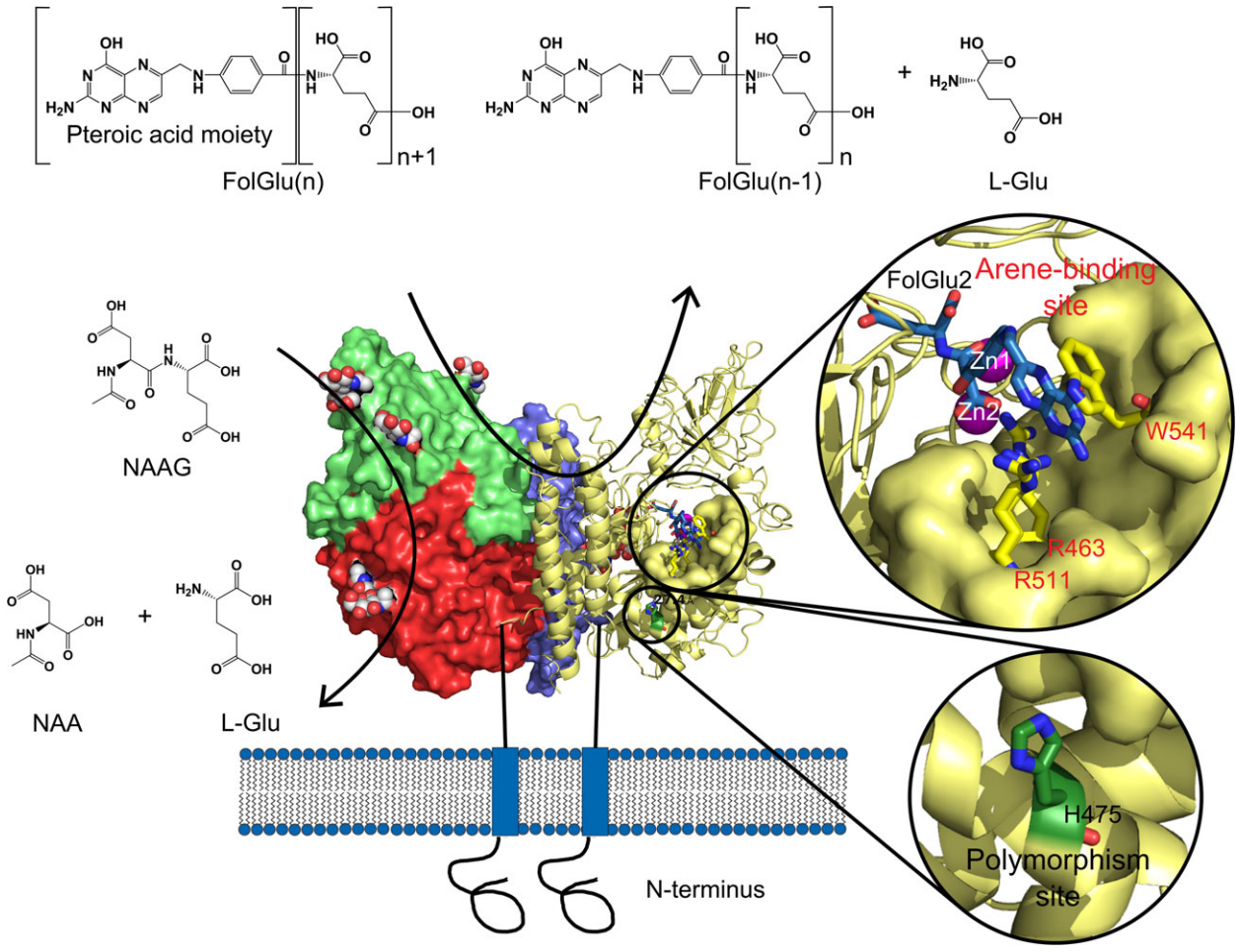

Fig. 1. Overall view of GCPII topology with the active site, ABS, His475Tyr mutation, and natural substrates NAAG and folyl-poly-γ-L-glutamates ($FolGlu_n$) highlighted. Note that folic acid (abbreviated as $FolGlu_0$) is pteroic acid linked to a single glutamate residue with a peptide bond. Mono-γ-L-glutamylated folic acid is abbreviated $FolGlu_1$. Pteroic acid is composed of a substituted pteridine moiety linked to p-aminobenzoic acid via a nitrogen atom. The left monomer of GCPII is shown in surface representation, with saccharide moieties depicted as gray spheres, the protease domain (amino acids 57–116 and 352–590) in red, the apical domain (amino acids 117–351) in green and the C-terminal domain (amino acids 591–750) in blue. The right monomer is shown in cartoon representation, with the ABS highlighted in the upper circle and the surface amino acid His475, subject of the disputed polymorphism, in the lower circle.

structural and enzymological aspects of $FolGlu_n$ cleavage in the small intestine have not been well-characterized. Numerous reports and epidemiologic studies on the naturally occurring His475Tyr polymorphism have analyzed the potential influence of this polymorphism on physiological levels of folates and folate-related metabolites (e.g. homocysteine). Some studies involving 600–1900 subjects concluded that this polymorphism influences neither folate, nor homocysteine levels [9–11]. By contrast, a study including 2700 subjects concluded that the His475Tyr polymorphism confers higher folate and lower homocysteine levels [8]. Other much smaller studies (30–44 participants) have correlated the polymorphism with altered folate levels and some pathologies, such as breast cancer [12]. Importantly, all of these studies investigated only genotype-phenotype association data, which makes their results rather inconclusive. For these reasons, the folyl poly-γ-L-glutamate hydrolyzing activity of GCPII and its natural His475Tyr variant, both at the structural and enzymological level, is the focus of the present study.

Structurally, the extracellular part of GCPII folds into three distinct domains: the protease domain (amino acids 57–116 and 352–590), the apical domain (amino acids 117–351) and the C-terminal domain (also called the dimerization domain; amino acids 591–750) (Fig. 1). All three GCPII domains contribute to formation of the substrate binding cavity, which consists of the S1 site, the active site and the S1′ site.

The active site features two Zn ions coordinated by the side chains of His377, Asp387, Asp453, Glu425 and His553. In the vicinity of the S1 site, there is a highly flexible segment called the 'entrance lid' (amino acids Trp541 to Gly548). In its closed conformation, the entrance lid fully shields the substrate binding cavity from the extramolecular space, as observed when small substrates such as NAAG are bound [4] [Protein Data Bank (PDB) code: 3BXM]. When the entrance lid is in open conformation, larger inhibitors can bind [13] (PDB code: 2XEF, 2XEG, 2XEI, 2XEJ). Similar to the S1 site, the S1′ site can be shielded from the extramolecular space by a flexible amino acid segment. The so-called 'glutarate-sensor' (Tyr692 to Ser704) takes on a β-stranded conformation upon binding of the glutamic acid moiety in the S1′ position, sealing off an otherwise present small funnel [6,14].

Recently, Zhang *et al.* [13] identified and structurally characterized an exosite of GCPII that binds aromatic moieties. This exosite, termed the arene-binding site (ABS), is formed by the first amino acid of the entrance lid, Trp541, together with Arg511 and Arg463. Adding a dinitrophenyl moiety with a length-optimized linker to a GCPII inhibitor significantly enhances the inhibitor's affinity toward GCPII via the avidity effect (i.e. by allowing it to bind to GCPII in a bidentate mode).

The present study aimed to determine: (a) how GCPII recognizes and cleaves polyglutamylated folates; (b) the role of the ABS in the recognition and binding of folyl-poly-γ-L-glutamic acids; and (c) whether the His475Tyr polymorphism alters the hydrolyzing activity of folyl-poly-γ-L-glutamate. Accordingly, we solved a series of X-ray structures of complexes between the inactive GCPII mutant, Glu424Ala, and several naturally occurring polyglutamylated folates ($FolGlu_{1–6}$). In addition, we analyzed the three-dimensional (3D) structure of the GCPII His475Tyr variant and compared its activity with the wild-type using a novel assay, and also performed structure–activity studies of the ABS.

Results

X-ray structures of recombinant human GCPII-Glu424Ala-$FolGlu_{1/2/3/4/5/6}$ complexes

To shed light on the binding mode of individual polyglutamylated folates in the substrate-binding pocket of GCPII, we determined X-ray structures of seven complexes of the hydrolytically inactive recombinant human GCPII (rhGCPII) Glu424Ala mutant [4] and $FolGlu_{1/2/3/4/5/6}$ substrates. The $FolGlu_{1–3}$ structures were refined at resolutions of between 1.65 and 2.00 Å, with good crystallographic parameters (Table 1). Their overall folds are almost identical. The only major differences were observed in the conformation of residues Arg463, Arg511 and Trp541 (forming the ABS) implicated in substrate binding. For $FolGlu_{1–2}$, the complete substrate molecules were defined in the electron density. For $FolGlu_3$, only the C-terminal di-γ-L-glutamyl-glutamate and pteridine moieties were defined in the electron density map (Fig. 2A). In the structures obtained using $FolGlu_{0/4/5/6}$ crystallization conditions, only di-γ-L-glutamyl-glutamate moieties could be modeled, and the remaining parts of these substrates were not defined in the electron density (data not shown).

As expected, the ultimate (C-terminal) glutamate, the scissile peptide bond and the penultimate (S1-bound) glutamate overlap in all seven structures (data not shown). The C-terminal glutamate is placed in the S1′ pocket in a manner identical to that described for previously reported structures [13,15], as indicated, for example, by the superposition of $FolGlu_{1–3}$ structures with the ARM-P4 structure (Fig. 2B). The carbonyl oxygen

Table 1. Data collection and refinement statistics for the X-ray models of rhGCPII-Glu424Ala-FolGlu$_{1–3}$ and rhGCPII-His475Tyr-Glu complexes (PDB code: 4MCP, 4MCQ, 4MCR and 4MCS, respectively). The number of ions includes two zinc, one calcium and one chlorine atom. Ramanchandran plot outliers are: Gly335, Val382 and Ser547 in the FolGlu$_1$ structure; Gly335 and Val382 in the FolGlu$_2$ structure; Val382 in the FolGlu$_3$ structure; and Gly335 and Val382 in the His475Tyr structure. All outliers were firmly defined by the electron density maps. Values in parentheses correspond to the highest-resolution shell.

	rhGCPII Glu424Ala-FolGlu1	rhGCPII Glu424Ala-FolGlu2	rhGCPII Glu424Ala-FolGlu3	rhGCPII His475Tyr-Glu
Data collection statistics				
Wavelength (A)	1.0000	0.91840	1.0000	0.91840
Temperature (K)	100	100	100	100
Space group	I222	I222	I222	I222
Unit cell parameters *a*, *b*, *c* (A)	101.5, 130.2, 158.8	101.6, 130.1, 159.9	101.5, 129.8, 158.8	101.5, 130.4, 159.0
Resolution limits	40.0–1.65 (1.71–1.65)	50.0–2.00 (2.07–2.00)	40.0–1.65 (1.71–1.65)	30.0–1.83 (1.90–1.83)
Number of unique reflections	119 318 (7644)	72 031 (6798)	114 830 (6220)	86 816 (8128)
Redundancy	6.9 (3.8)	5.0 (4.4)	7.0 (4.1)	4.0 (2.9)
Completeness	94.5 (61.1)	99.3 (94.8)	91.1 (49.9)	93.3 (88.3)
$I/\sigma(I)$	22.6 (2.8)	30.1 (5.0)	24.2 (3.1)	16.5 (2.1)
R_{merge}	0.077 (0.320)	0.051 (0.226)	0.076 (0.299)	0.067 (0.429)
Refinement statistics				
Resolution limits (A)	30.0–1.65 (1.69–1.65)	22.96–2.00 (2.05–2.00)	29.5–1.65 (1.69–1.65)	28.5–1.83 (1.88–1.83)
Total number of reflections	117 931 (5137)	69 042 (4495)	113 362 (4236)	84 871 (5872)
Number of reflections in the working set	116 731 (5078)	67 854 (4431)	112 222 (4185)	83 135 (5750)
Number of reflections in the test set	1200 (59)	1188 (64)	1140 (51)	1736 (122)
R/R_{free}	0.150/0.164 (0.220/0.265)	0.137/0.172 (0.146/0.206)	0.133/0.167 (0.221/0.260)	0.159/0.191 (0.251/0.284)
Total number of nonhydrogen atoms	6405	6267	6365	6384
Number of ligand atoms	41	50	59	19
Number of ions	4	4	4	4
Number of water molecules	668	521	610	584
Average B–factor (A2)				
Protein atoms	28.8	27.1	26.1	30.7
Water molecules	36.0	31.3	32.1	39.6
Substrate	38.1	40.6	67.0	31.1
Wilson *B*-factor	29.6	27.6	26.1	31.5
r.m.s.d.				
Bond lengths (A)	0.015	0.021	0.011	0.019
Bond angles (°)	1.49	1.66	1.38	1.61
Planarity (A)	0.008	0.008	0.007	0.009
Chiral centers (A3)	0.110	0.118	0.092	0.120
Ramanchandran plot (%)				
Favored	97.1	97.2	97.5	97.4
Poor rotamers	1.36	1.02	1.19	1.36
Outliers	0.44	0.29	0.15	0.29
Gaps in the structure	44–55, 654–655	44–55, 654–655	44–55, 654–655	44–54, 541–543, 654–655

of the scissile peptide bond is polarized by the Zn1 ion (2.5 Å; not visible in Fig. 2B) and the polarization/positioning is further assisted by its interactions with the Tyr552 hydroxyl group (2.7 Å) and His553 Nε2 (3.1 Å). The peptide amide group donates a hydrogen bond to the main-chain carbonyl oxygen of Gly518 (3.0 Å). The α-carboxylate of the penultimate glutamate directly engages positively charged Arg534 (2.6 Å) and Arg536 (3.1 Å). The above-mentioned interactions are structurally conserved for both folate-based and NAAG substrates [4].

More structural diversity was observed for the distal (pteroate) parts of individual substrates. In the case of FolGlu$_1$ (Fig. 2A), the 'single glutamate linker' is too

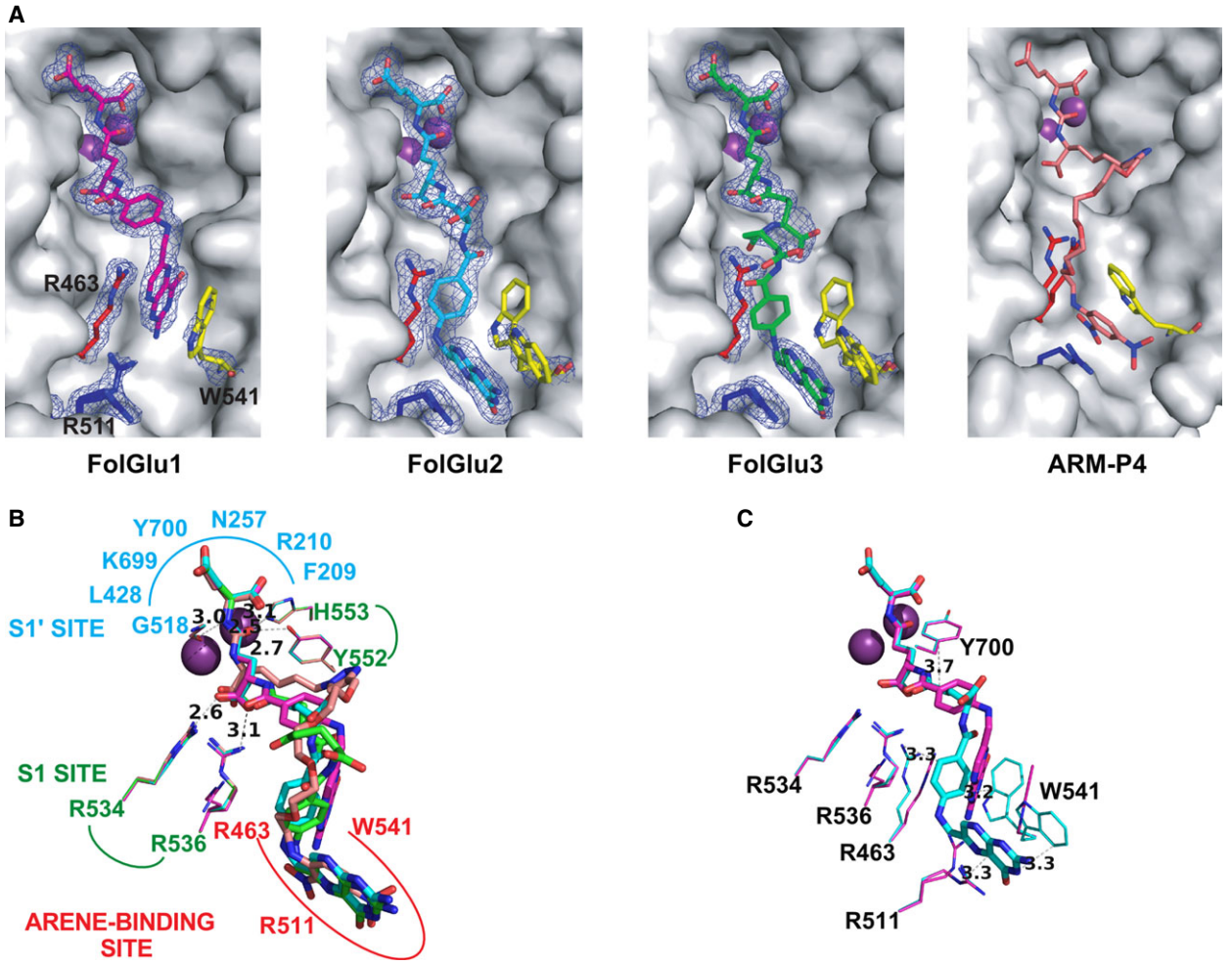

Fig. 2. Binding modes of $FolGlu_{1/2/3}$ and ARM-P4 (PDB code: 4MCP, 4MCQ, 4MCR and 2XEG, respectively) to rhGCPII Glu424Ala. (A) $FolGlu_{1–3}$ substrates and the ARM-P4 inhibitor are shown in stick representation, whereas the selected part of GCPII is shown in surface representation. Residues forming the ABS are shown in stick representation and are colored red (Arg463), blue (Arg511) and yellow (Trp541). Note the similar positioning of the distal pteroate ($FolGlu_{2/3}$) and dinitrophenyl (ARM-P4) functionalities in the ABS. By contrast, the glutamylation status of $FolGlu_1$ is too low to allow the distal pteroate moiety to fully engage the ABS. $FolGlu_{1/2/3}$ substrates and ABS residues are shown with their corresponding $2F_o - F_c$ electron density maps contoured at 1σ. Note the weak electron density maps of the Trp541 residue, reflecting its inherent flexibility (and allowing to model only one conformation for $FolGlu_1$ and two conformations for $FolGlu_{2–3}$). Similarly, when $FolGlu_3$ substrate is bound, a weak electron density map is observed for the p-aminobenzoic acid and the most distal glutamic acid. (B) Superposition of $FolGlu_{1/2/3}$ and ARM-P4. Substrates/inhibitor are shown in stick representation with carbon atoms colored magenta ($FolGlu_1$), cyan ($FolGlu_2$), green ($FolGlu_3$) and pink (ARM-P4). The active-site zinc atoms are shown as purple spheres. For clarity, amino acids of the S1′ site, S1 site and ABS are indicated only by their names. Note the structural overlap among the ultimate and penultimate glutamates and the variability of the distal substrate/inhibitor parts. (C) Superposition of $FolGlu_1$ (magenta) and $FolGlu_2$ (cyan) structures. Substrates are in stick representation; selected GCPII residues are shown as lines. Note the flexibility of the ABS residues (Arg463, Arg511 and Trp541), which is crucial for engagement of the distal pteroate moiety of the substrates by GCPII, whereas the S1 site residues (Arg 534, Arg536, Tyr700) take up almost the same conformation in both structures. The active-site zinc atoms are shown as purple spheres; selected hydrogen bonds are shown as broken lines.

short for the benzoate/pteroate moieties to reach fully into the ABS. Instead, the p-amino benzoyl moiety of $FolGlu_1$ is engaged in a staggered π-π stacking interaction with the S1′ and S1 side chain of Tyr700 (3.7 Å) (Fig. 2C), whereas the pteridine ring system interacts in a π-cation fashion with the side chain of Arg463 (Fig. 2A). Another consequence of the substrate being too short is that the electron desity map allows the Trp541 to be modeled in only one conformation (whereas, for $FolGlu_{2–3}$ complexes, two conformations are possible). The Trp541 residue is in a conformation of π-π stacking with the pteroate moiety. For $FolGlu_2$

(Fig. 2A), which features a one-Glu-residue-longer linker, π-cation and T-shaped stacking interactions were observed between the p-aminobenzoyl group of the substrate and side chains Arg463 (3.3 Å) and Trp541 (3.2 Å), respectively (Fig. 2C). The pteridine ring is wedged between the side chains of Arg511 and Trp541 with its plane parallel to the indole of Trp541 (3.3 Å) and the guanidinium group of Arg511 (3.3 Å). We observed two possible conformations for Trp541, as well as higher B-factors for its side chain. A simple overlay of $FolGlu_1$ and $FolGlu_2$ structures is provided in Fig. 2C, with amino acids of the ABS (Arg463, Arg511, Trp541) and the S1 site (Arg 534, Arg536, Tyr700) shown as lines. This overlay illustrates that the S1 site residues (Arg 534, Arg536 and Tyr700) are essentially in the same conformation (except for a double conformation of Arg536 in the $FolGlu_1$ structure), whereas the ABS residues (Arg463, Arg511, Trp541) adopt different conformations according to the length of the γ-glutamyl chain of the folic acid. For $FolGlu_3$ (Fig. 2A), the binding mode and positioning of the pteroate moiety are almost identical to those observed for $FolGlu_2$, with Trp541 present in two conformations. However, the $2F_o - F_c$ electron density for the distal parts of $FolGlu_3$ was much weaker, and the B-factors of the final model were much higher compared to those calculated for the structure of the $FolGlu_2$ complex. As already noted above, $FolGlu_3$ substrate has only the C-terminal di-γ-L-glutamyl-glutamate and pteridine moieties defined in the electron density map, and so the p-aminobenzoic acid moiety and the most distal glutamic acid residue are modeled rather only sterically. These factors highlight the conformational flexibility of $FolGlu_3$ and the limited contribution of its distal part to GCPII binding. In general, the side chain of Trp541 was defined by a weak $2F_o - F_c$ electron density in our structures (Fig. 2A) and the variety of conformations observed indicates the inherent flexibility of this residue. In summary, the structures of $FolGlu_{1–6}$ complexes reported in the present study clearly suggest that His475 is a surface residue too distant from the binding cavity and cannot play a role in the interaction of GCPII with folyl-poly-γ-L-glutamic acids.

Structural characterization of recombinant human GCPII His475Tyr

To determine the putative effects of the His475Tyr substitution on the 3D structure of GCPII and on the enzyme's substrate binding and processing, we co-crystallized the rhGCPII His475Tyr variant with the natural substrate NAAG. Similar to our previous observations [3] for the wild-type enzyme, glutamate (the product of the enzymatic reaction) was observed in the S1′ site of the final model, which was refined at 1.83 Å resolution, with crystallographic *R*-factors equal to 0.159 (R_{free} = 0.191; for complete data collection and refinement statistics, see Table 1).

The His475Tyr/glutamate structure (PDB code: 4MCS) was analyzed and compared with the corresponding complex of glutamate with the wild-type enzyme [3] (PDB code: 2C6G). The overall fold is almost identical for both the His475Tyr and wild-type proteins, as illustrated by a r.m.s.d. of 0.19 Å (for the 661 equivalent Cα pairs). Additionally, no differences were observed in the arrangement of the internal substrate-binding cavity and residues in the active site or in the positioning and binding mode of glutamate. Most importantly, both proteins share an almost identical arrangement of amino acids in the vicinity of residue 475, and the imidazole ring of wild-type His475 and the benzene ring of Tyr475 overlap spatially (Fig. 3).

Site-directed mutagenesis study of the ABS

The structural data reported in the present study indicate that at least some $FolGlu_n$ substrates interact with GCPII in a bidentate mode: the C-terminal and penultimate glutamates engage residues of the S1′ and S1 sites, respectively, whereas the pteridine double-ring and p-aminobenzoyl moiety interact with the residues of the recently identified ABS [13]. To investigate the role of the ABS in folate binding and hydrolysis, a series of GCPII variants with mutations in the ABS (Arg463Leu, Arg511Leu and Trp541Ala) was prepared by site-directed mutagenesis, expressed in *Drosophila* S2 cells, and purified to homogeneity. The activities of these enzymes were characterized in terms of their hydrolysis of NAAG and $FolGlu_4$ (substrates with structurally no and a poorly defined interaction with the ABS, respectively) and $FolGlu_{1–3}$ (substrates with structurally well-defined interactions of the pteridine double-ring with the ABS). We assumed that, if interactions between $FolGlu_{1–3}$ and the ABS influence the kinetic parameters of substrate hydrolysis, a decrease in substrate affinity (i.e. an increase in K_M value) should be observed for the GCPII ABS mutants. Although no major differences between the three mutants and the wild-type enzyme were observed for the kinetic parameters of $FolGlu_1$ hydrolysis, there was a slight decrease in K_M values for both NAAG and $FolGlu_{2–4}$ (Fig. 4). However, because the binding of NAAG does not engage residues of the ABS [13], the observed changes cannot be linked to the interaction of the pteroate with the ABS.

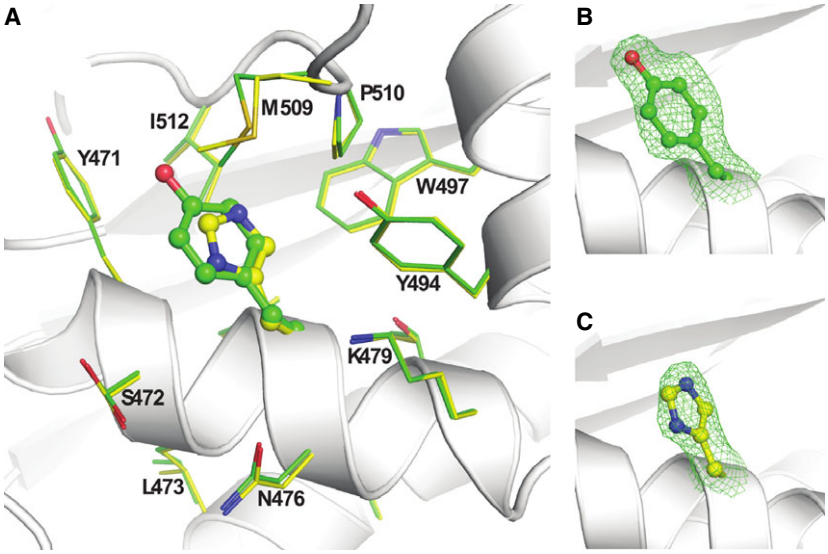

Fig. 3. Superposition of wild-type GCPII (PDB code: 2C6G) and His475Tyr variant (PDB code: 4MCS). (A) Detailed view of amino acids in the vicinity (5.0 Å) of the polymorphism site. Atoms are colored according to element, with carbons colored green and yellow for the wild-type and His475 variants, respectively. His475 and Tyr475 are depicted in ball-and-stick representation; surrounding residues (within 5 Å) are shown as lines. The main chain (cartoon in gray) is from the His475Tyr structure. (B,C) Detailed images of the Tyr475 (B) and His475 (C) residues showing the $F_o - F_c$ electron density contoured at 3.0 σ.

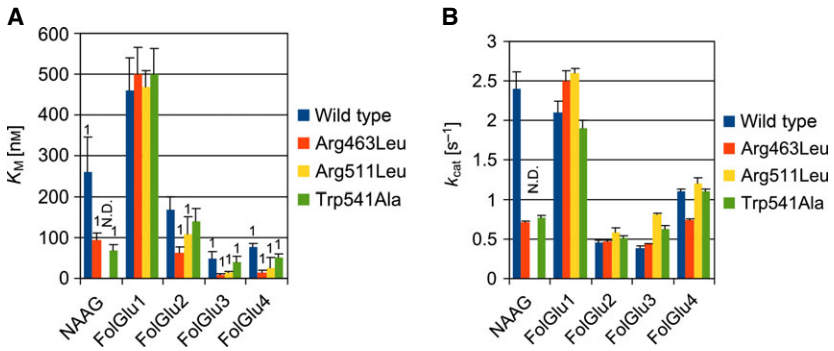

Fig. 4. Comparison of the kinetic parameters of $FolGlu_n$ hydrolysis by GCPII wild-type and ABS mutant variants. This comparison shows differences between wild-type and the mutant variants Arg463Leu, Arg511Leu and Trp541Ala in K_M (A) and k_{cat} (B). The reaction buffer was 25 mM Tris (pH 7.5). Error bars indicate the SD derived from a simple hyperbolic fit of a single saturation curve. '1' indicates extrapolation of K_M (i.e. K_M value lower than the lowest substrate concentration). N.D., not determined.

To analyze the relevance of the ABS for interactions with small molecule ligands, we compared the inhibition profiles of four GCPII-specific inhibitors [13] (Fig. 5) towards the Trp541Ala mutant and the wild-type enzyme. Three of these inhibitors (ARM-P2, ARM-P4 and ARM-P8; urea based inhibitors containing the ABS-targeted dinitrophenyl moiety linked via 2/4/8 ethyleneglycol units) have been reported to engage the ABS. On the other hand, the MeO-P4 compound (urea based inhibitor with four ethyleneglycol units containing no arene-binding site targeted moiety) is incapable of any interactions with ABS residues [13]. The inhibition data are summarized in Fig. 5. As expected, we observed no difference between the mutant and wild-type proteins in the K_i values for MeO-P4. By contrast, two inhibitors featuring the distal dinitrophenyl moiety linked with an optimal linker (and thus capable of ABS interactions; ARM-P4, ARM-P8) demonstrate an affinity that was several-fold higher for the wild-type protein compared to Trp541Ala (Fig. 5B), thus confirming the relevance of the ABS (namely Trp541) for the binding of aromatic moieties. Moreover, for ARM-P4, we observed an approximately one-order-of-magnitude difference in K_i

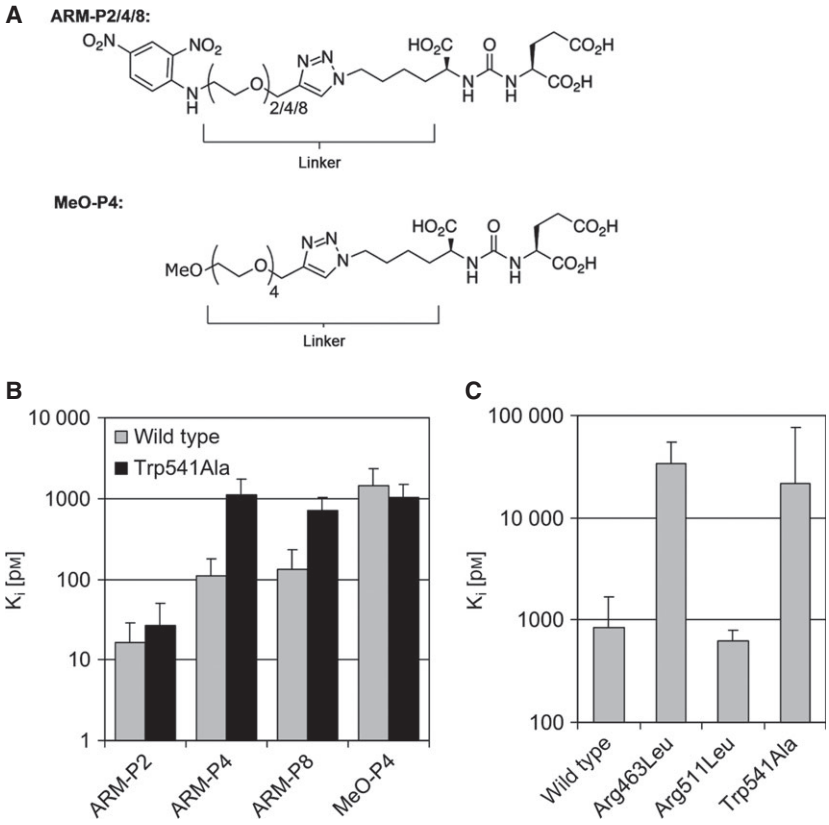

Fig. 5. Inhibition studies with inhibitors ARM-P2, ARM-P4, ARM-P8 and MeO-P4. (A) Structures of the inhibitors ARM-P2, ARM-P4, ARM-P8 and MeO-P4. Inhibitor structures are adapted with permission [13]. (B) Comparison of K_i values determined for inhibitors ARM-P2, ARM-P4, ARM-P8 and MeO-P4 towards GCPII wild-type and Trp541Ala mutant with $FolGlu_1$ as a substrate. Trp541 is an important element for interacting with ABS-targeted aromatic moieties. For ARM-P2, ARM-P8 and MeO-P4, the $FolGlu_1$ substrate concentration was 80 nM, whereas, for ARM-P4, it was 160 nM. (C) K_i values for ARM-P4 toward GCPII wild-type and Arg463Leu, Arg511Leu and Trp541Ala mutants. Error bars indicate the SD derived from a sigmoidal fit of a single inhibition curve.

values for Trp541Ala and Arg463Leu compared to wild-type (Fig. 5C; for the structure of GCPII in complex with ARM-P4, see Fig. 2A).

Enzymatic characterization of recombinant human GCPII wild-type and His475Tyr

A single study describes a ~ 50% decrease in the folate-hydrolyzing activity of the His475Tyr variant compared to the wild-type enzyme [16]. To provide a detailed enzymatic characterization and side-by-side comparison of the wild-type and His475Tyr variants, we developed a novel UPLC-based assay and measured the kinetic parameters of hydrolysis of a panel of natural GCPII substrates by purified recombinant proteins (Fig. 6). Overall, K_M and k_{cat} values for hydrolysis of the substrates by wild-type and His475Tyr GCPII were very similar. However, we observed an almost one-order-of-magnitude difference in affinity (difference in K_M values) for both enzymes for $FolGlu_n$ substrates with more than one C-terminal glutamates (compare $FolGlu_1$ and $FolGlu_{2–6}$ substrates; Fig. 6A). This clearly suggests additional interaction(s) between the $FolGlu_{2–6}$ substrates and the protein or, alternatively, lesser steric hindrance. Somewhat unexpectedly, we found differences between the wild-type

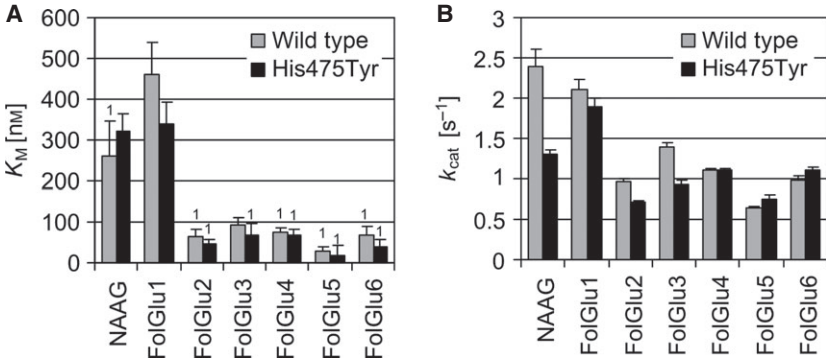

Fig. 6. Comparison of the kinetic parameters of rhGCPII wild-type and rhGCPII His475Tyr for hydrolysis of a panel of natural substrates (NAAG and $FolGlu_{1–6}$). This comparison shows differences between the wild-type and His475Tyr variant in K_M (A) and k_{cat} (B). Error bars indicate the SD derived from a simple hyperbolic fit of a single saturation curve. '1' indicates extrapolation (i.e. K_M value lower than the lowest substrate concentration).

and His475Tyr variants when NAAG was used as a substrate (approximately the same K_M but with a different k_{cat}; Fig. 6B); however, this difference was not statistically significant.

Thermal stability of recombinant human GCPII wild-type and His475Tyr

Although the wild-type and His475Tyr variants are almost indistinguishable at the structural level, we investigated the possibility that the His to Tyr substitution influences the thermal stability of the His475Tyr variant. Accordingly, we used a Thermofluor assay to define the temperature midpoint for the unfolding transition, T_m. The T_m value determined for the wild-type protein (67.2 ± 0.118 °C) is comparable to that determined for the His475Tyr variant (69.2 ± 0 °C), in disagreement with a biologically significant (de)stabilization effect of the mutation.

The kinetic parameters obtained for folate-based substrates, together with the data obtained by 3D structure analysis and thermal stability experiments, suggest that the His475Tyr variant is equivalent to the wild-type enzyme at both the structural and enzymological levels.

Discussion

The present study aimed to investigate the molecular recognition of polyglutamylated folates by GCPII and to clarify whether the reported His475Tyr polymorphism might alter the binding and/or turnover of polyglutamylated folates, potentially leading to altered levels of folates and related metabolites in humans. Furthermore, we set out to analyze the role of the ABS, a recently identified putative exosite in the GCPII structure, with respect to the molecular recognition of polyglutamylated folates and, more generally, in the binding of ligands containing an aromatic moiety capable of interacting with this exosite.

Although extensive epidemiologic studies on the effect of the His475Tyr polymorphism have been conducted, the possible role of this mutation in the folate uptake and metabolism of metabolically related molecules remains unclear. Folates are essential for the C1 metabolism of cells (i.e. for the biosynthesis of nucleobases) and they are important for the replication of rapidly dividing cancer cells. Folate insufficiency might result in altered levels of folate-related metabolites and/or in an altered susceptibility to certain types of cancer. However, it is difficult to speculate how a single amino acid mutation in a region distant from the active site (Fig. 1; His475 is located 27 Å from the active site Zn1) could influence the catalytic activity of the enzyme. Nevertheless, there have been reports of long-range rearrangements of an enzyme's substrate binding cleft caused by very distant mutations; for example, in HIV-1 protease [17]. Thus, the idea that the His475Tyr mutation could influence GCPII activity was worthy of consideration. However, we found that the enzymatic properties of recombinant human GCPII His475Tyr are very similar to those of the wild-type enzyme (almost identical K_M values and turnover numbers toward all tested folyl-poly-γ-L-glutamate substrates). Furthermore, 3D structure analysis of the protein by X-ray crystallography did not reveal any significant structural changes caused by the mutation. The very subtle differences that we observed cannot account for the reported phenotype of the polymorphism [8–12].

It is relevant to note that all recombinant proteins discussed in the present study were prepared in *Drosophila* S2 cells and comprise the full ectodomain of human GCPII in which the enzymatic activity resides. We have previously shown that the enzymatic activity of this recombinant protein is indistinguishable from that of native, full-length GCPII [18].

All seven structures solved as part of the present study provide structural explanation as to why the His475Tyr mutant behaves almost identically to the wild-type enzyme. Our structural analysis showed that the bound $FolGlu_{1–6}$ substrates do not interact with His475. Additionally, the structure of GCPII His475Tyr exhibits no significant re-arrangement of the substrate binding pocket. Although the enzymologic comparison of rhGCPII wild-type and His475Tyr identified no significant differences in the cleavage of $FolGlu_n$, thermal stability, or structure, the possibility remains that the His475Tyr mutation causes altered folic acid levels and folate-related metabolites in humans. The His475Tyr mutation may influence folate levels by other mechanisms (e.g. by changing the trafficking of GCPII, by changing its expression level or by changing the enzyme's interaction with potential signaling partners). These potential mechanisms of action remain to be investigated.

There was one peculiarity about the X-ray structure of the complex obtained from the crystallization condition with the inactive GCPII mutant (Glu424Ala) and $FolGlu_0$. As already noted, this structure contains a di-γ-L-glutamyl-glutamic moiety, even though no compound containing such a moiety has been added to the crystallization drop. A possible explanation is that some higher polyglutamylated folates ($FolGlu_{>4}$) originating from the cultivation medium may have remained bound throughout the purification process

(only a one-step affinity purification), not allowing the $FolGlu_0$ to bind. In line with such a hypothesis, folic acid did not inhibit the cleavage of $FolGlu_3$ (data not shown), suggesting that the folate (the product) has a lower affinity to the enzyme than polyglutamylated folate (the substrate).

Inspection of the kinetic properties of the wild-type and His475Tyr variants revealed that folate substrates harboring more than two C-terminal glutamates bound to the enzyme more effectively (by almost one order of magnitude), as reflected by differences in their K_M values (compare $FolGlu_1$ and $FolGlu_{2–6}$; Fig. 6A). This difference in K_M is in good agreement with the previously published observation that the di-γ-L-glutamate moiety inhibited the NAAG-hydrolyzing activity of GCPII twenty-fold less effectively than the tri-γ-L-glutamate moiety, whereas tri-, tetra- and penta-γ-L-glutamate moieties were observed to inhibit to a similar extent [19].

To further analyze the contribution of the ABS to substrate recognition, we performed site-directed mutagenesis of the key residues of the exosite (Arg463, Arg511 and Trp541). Kinetic analysis of the recombinant mutant proteins did not show a major influence of the ABS residues on polyglutamyl-folate binding and turnover, and only a modest, approximately three-fold increase in NAAG binding affinity. Because NAAG cannot engage the ABS, the direct effects of these mutations on substrate recognition can be excluded and more intricate/subtle (unidentified) contributions (e.g. the flexibility of the entrance lid) may play a role.

The binding of selected inhibitors targeting the ABS, on the other hand, was shown to be influenced by their interaction with the residues forming the ABS both structurally and by using inhibition studies [13]. The present study tested those inhibitors with rhGCPII wild-type and Trp541Ala variants [13] (structures of the inhibitors are shown in Fig. 5A). Inhibitors that had been previously reported to interact with the ABS showed a significant loss of binding to the GCPII Trp541Ala variant, in which the ABS is presumably disrupted (Fig. 5B), confirming the importance of this structural feature for inhibitor binding. It is somewhat unexpected that, on the one hand, the crystallographic and site-directed mutagenesis data of the ABS provide a good mechanistic explanation for changes in inhibitor affinity, whereas, on the other hand, there is a disconnection between the structural data (ABS engages the pteroate) and kinetic studies (mutations in the ABS mostly do not change the kinetic parameters of substrate hydrolysis). We can speculate that these differences might be attributed to differences in the kinetics of substrate versus inhibitor interactions with GCPII. The crystallographic data represent the most preferred conformation of a given GCPII complex achieved in a time-scale of several days (crystal growth). In addition, when a typical inhibition constant is determined, the inhibitor is usually preincubated with the enzyme for several minutes (5–13 min in this case), giving the inhibitor sufficient time to adopt the most energetically favorable conformation. On the other hand, a $k_{cat} > 1\ s^{-1}$ indicated that a given substrate is bound and hydrolyzed in less than a second. Although the high-affinity C-terminal part of the substrate is docked into (and released from) the active site of GCPII within this time scale, the flexible distal part might not be able to engage the ABS and thus has limited (or no) contribution to the overall affinity. This conjecture is supported by binding kinetics observed for GCPII interactions with a fluorescently labeled GCPII inhibitor/probe [20]. We have observed that the BODIPY fluorophore engages the ABS (C. Bařinka, unpublished work) and that this engagement leads to quenching of fluorescence intensity. The fluorescence quenching is rather gradual, taking tens of seconds, rather than being instantaneous. This suggests a bidentate mode of substrate/inhibitor binding, in which the C-terminal (docking) part of the inhibitor binds 'immediately' and the distal part binds somewhat more slowly. Another possibility could be that this exosite is inherently a lower-affinity site allowing the binding of all forms of folic acid moiety (folyl/5-methyl-dihydrofolyl/5-methyl-tetrahydrofolyl). The natural diet contains predominantly 5-methyl-tetrahydrofolyl-poly-γ-L-glutamates [21], which are less stable than their fully oxidized folyl-forms and may easily be oxidized to 5-methyl-dihydrofolyl-poly-γ-L-glutamates or may even degrade to the folyl-forms. The same may apply to the folate receptor α isoform, which has been recently co-crystallized with the folic acid but not with 5-methyl-tetrahydrofolic acid [22] (PDB code: 4LRH), although this isoform of folate receptor is expressed in kidney in renal tubular cells to resorb 5-methyl-tetrahydrofolic acid, a major form of folic acid found in blood.

Another indication that the ABS is a rather low-affinity binding site is provided by the x-ray structure of GCPII in complex with a nonhydrolyzable methotrexate analogue of the $FolGlu_1$ substrate (PDB code: 3BI1) [23]. In this structure, only the γ-D-glutamyl-L-glutamate moiety with a spatial arrangement analogous to that seen in $FolGlu_{1–3}$ structures was defined by the electron density map, whereas the pteridine ring was disordered and no stacking with Trp541 could be observed [23].

Measurements of kinetic parameters also revealed that the mechanism by which polyglutamylated folates

are cleaved is probably distributive (i.e. the γ-linked L-glutamates from folyl-poly-γ-L-glutamic acids are cleaved sequentially from the C-terminus). For higher $FolGlu_n$ turnovers (10–20%), we have always observed a small amount (~ 1–2%) of the $FolGlu_{n-2}$ product (data not shown; the contribution of the $FolGlu_{n-2}$ product to the kinetic data could thus be neglected).

The data reported in the present study might be a starting point for the design of a homolog-specific ligand or inhibitor of GCPII and GCPIII. Although these two homologous proteins share 81% similarity at the amino acid sequence level [24] and their 3D structures are very similar [25], some studies suggest that GCPIII might play a distinct enzymatic role in specific tissues, such as the testes [26]. As shown by Zhang *et al.* [13], the choice of a proper ABS-targeting moiety can improve inhibitor potency by several orders of magnitude. Because GCPIII does not appear to have a corresponding ABS in its structure, an appropriate inhibitor targeting this exosite in GCPII might represent a useful tool for distinguishing both enzymes at the protein level.

Materials and methods

Preparation of rhGCPII

rhGCPII wild-type and mutants were heterologously overexpressed in *Drosophila* Schneider 2 (S2) cells [27] using the pMT/Bip/AviTEV/rhGCPII vector, as described previously [28]. In this expression system, the soluble extracellular part of GCPII N-terminally fused with a so-called AviTEV tag is secreted into the medium. The AviTEV tag comprises an *in vivo* biotinylated Bir ligase recognition sequence (AviTag™; Avidity, Aurora, CO, USA) separated from the target protein with a tobacco etch virus (TEV) recognition sequence. Proteins were purified using the Streptavidin Mutein Matrix™ (Roche Diagnostics, Basel, Switzerland) [28].

Site-directed mutagenesis and cloning

The Glu424Ala and His475Tyr mutations were introduced by cutting the coding sequence out of the corresponding pMTNAEXST vector [4] using the restriction enzymes *Bgl*II/*Xho*I (New England Biolabs, Ipswitch, MA, USA) and *Xho*I/*Xcm*I (New England Biolabs), respectively, and re-ligating into the pMT/Bip/AviTEV/rhGCPII vector using T4 DNA ligase (New England Biolabs). The mutations Arg463Leu, Arg511Leu and Trp541Ala were introduced by mutating the pMT/Bip/AviTEV/rhGCPII vector using the QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA). Nucleotide sequences (5′ to 3′) of the primers were: GAAGGAAACTACACTCTGTTAGTTGATT GTACACCGCTGATG (forward primer, Arg463Leu), CATCAGCGGTGTACAATCAACTAACAGAGTGTAG TTTCCTTC (reverse primer, Arg463Leu), GTTCAGT GGCATGCCCCTGATAAGCAAATTG (forward primer, Arg511Leu), CAATTTGCTTATCAGGGGCATGCCACT GAAC (reverse primer, Arg511Leu), GTATACTAAA AATGCCGAAACAAACAAATTCAG (forward primer, Trp541Ala) and CTGAATTTGTTTGTTTCGGCATTT TTAGTATAC (reverse primer, Trp541Ala). Before synthesis, all primer sequences were checked using OLIGOCALC [29]. Sequences of the resulting plasmids were verified by dideoxynucleotide termination sequencing. Yields of pure recombinant proteins were 0.63 mg·100 mL^{-1} of media for the wild-type enzyme [28], 83 μg·100 mL^{-1} for His475Tyr, 9.5 μg·100 mL^{-1} for Arg463Leu, 6.3 μg·100 mL^{-1} for Arg511Leu and 0.10 mg·100 mL^{-1} for Trp541Ala.

Crystallization and data collection

Crystals were grown by the hanging drop vaporization technique at 291 K, as described previously [30], with minor modifications. The crystallization drops consisted of 1 μL of rhGCPII Glu424Ala mixed with an equal volume of the reservoir solution [33% (v/v) pentaerythritol propoxylate PO/OH 5/4 (Hampton Research, Aliso Viejo, CA, USA), 0.5% (w/v) poly(ethylene glycol) 3350 (Sigma-Aldrich, St Louis, MO, USA), 0.10 M Tris-HCl (Promega, Madison, WI, USA), pH 8.0] pre-mixed with 10 mM folyl-(0/1/2/3/4/5/6)-γ-L-glutamic acid (Schircks Laboratories, Jona, Switzerland) in a 9 : 1 (v/v) ratio. The His475Tyr variant was co-crystallized with NAAG at a final concentration of 2 mM in the same set-up. Crystals usually appeared within 1 day and grew to a final size of ~ 0.40 × 0.40 × 0.2 mm during the next few few weeks. Diffraction data were collected at 100 K using synchrotron radiation at the SER-CAT ID/22-BM beamlines (Argonne, IL, USA; 1.00 Å; MARMOSAIC 225; complexes with $FolGlu_{1/3/4}$) and at the MX 14.2 beamline (BESSYII, Helmholtz-Zentrum, Berlin, Germany; 0.9184 Å; MX-225; the His475Tyr variant and complexes with $FolGlu_{2/5/6}$). Complete datasets were collected from single crystals, and data were processed using either HKL2000 [31] or XDSAPP [32]. The complex with $FolGlu_2$ was tested in-house at 120 K on the MAR345 detector (MAR Research, Hamburg, Germany) using Cu Kα wavelength (1.5418 Å).

Structure determination and refinement

Structures of GCPII Glu424Ala complexes were determined by molecular replacement method using the coordinates of the GCPII-Glu424Ala-NAAG complex (PDB code: 3BXM) as a starting model [4]. Refinement calculations

and manual rebuilding were performed with REFMAC, version 5.5 [33] and COOT, version 6.1 [34], respectively. Approximately 1.0–1.7% of the data (corresponding to 1140–1736 reflections) were excluded from calculations and used instead to calculate the R_{free} value (R_{free} sets were independent). Models for individual substrates together with associated restraints were prepared using the PRODRG server [35] and were placed into the positive $F_o - F_c$ electron density maps, located in the substrate binding cavity of rhGCPII. An isotropic refinement protocol was used throughout all stages of the refinement. The stereochemical quality of the final models was evaluated using MOLPROBITY [36]. Data collection and structural refinement statistics are shown in Table 1.

GCPII activity assay

Reactions with the folyl-(1/2/3/4/5/6)-γ-L-glutamic acid substrates (Schircks Laboratories) and NAAG (Sigma-Aldrich) were performed in 25 mM Tris-HCl (pH 7.5) in a total volume of 250 μL. Concentrations of substrate stock solutions were determined by amino acid analysis. All components were pipetted at 0 °C, and reactions were started by adding enzyme and mixing. The reactions were placed in a 37 °C thermo-block (MB-102, Bioer, Hangzhou, China, or Thermomixer Comfort, Eppendorf, Hamburg, Germany) for 20 min. These reaction conditions were designed to yield 10–20% conversion of substrate. Reactions were stopped by adding 3.5 μL of stopping solution [72 μM 2-(phosphonomethyl)-pentanedioic acid (2-PMPA), 7.2 mM 2-mercaptoethanol (Sigma-Aldrich), 25 mM Tris, 0.48 M phosphoric acid (Penta, Praha, Czech Republic)]. Typically, nine data points were plotted for each substrate (with each point representing the mean of duplicate experiments) and saturation curves were fitted using GRAFIT [37].

For inhibition and kinetic studies with $FolGlu_{1–6}$ as a substrate, reactions were performed in 25 mM bistrispropane-HCl (Sigma-Aldrich) (pH 7.5). Reactions were carried out either in a total volume of 215 μL in a 96-well plate immersed in a water bath (Grant Instruments, Shepreth, UK) or in a total volume of 250 μL in 1.5-mL test tubes in a thermo-block (MB-102). All components except the substrate were pipetted at 0 °C. After a 5–13 min of preincubation at 37.0 °C (5 min of preincubation in the case of the 1.5-mL test tubes), substrate was added. The reaction proceeded for 20 min and was terminated by the addition of 10.0 μL of stopping solution [0.44 M phosphoric acid and 21.3 μM 2-PMPA (96-well plate) or 3.48 μL of 1.5 M phosphoric acid and 71.3 μM 2-PMPA (1.5-mL test tubes)]. The identity of the inhibitors was verified by Q-TOF Micro (Waters, Milford, MA, USA) and LTQ Orbitrap XL (Thermo Fisher Scientific, Waltham, MA, USA). Each inhibition curve was fitted using GRAFIT [37] into at least 12 data points typically acquired in duplicate.

Novel UPLC method for assaying glutamate carboxypeptidase activity using folyl-(1/2/3/4/5/6)-γ-L-glutamic acids as substrates

$FolGlu_n$ reaction mixtures were analyzed on a C18 Acquity UPLC HSS T3 2.1 x 100 mm column with 1.8-μm particles (Waters), guarded by a 0.22-μm pre-filter (Waters) and VanGuard pre-column (Waters), coupled to an Agilent 1200 Series UPLC instrument (Agilent Technologies, Santa Clara, CA, USA). Mobile phase A was 25 mM sodium phosphate buffer (pH 6.0) (Penta), supplemented with 0.02% (w/v) sodium azide (Sigma-Aldrich); mobile phase B was acetonitrile (VWR International, Radnor, PA, USA). Elution of individual $FolGlu_{1/2/3/4/5/6}$ molecules and their cleavage products was performed isocratically at 2.0%/1.5%/1.1%/0.4%/0.2%/0.0% acetonitrile, respectively. The column temperature was set to 50.0 °C. The HPLC runs consisted of 1.8 min of isocratic flow at 0.0–2.0% B, 0.1 min of transition to 10.0% B, 1.1 min at 10.0% B, 0.1 min of transition back to 0.0–2.0% B and 7.4 min of re-equilibration. Analytes were detected at 281 nm and 354 nm. Inhibition reactions with $FolGlu_1$ were analyzed in the same manner, except that the percentage of acetonitrile was 2.1% instead of 2.0%. The substrate turnover was quantified as the ratio of the substrate and product peak areas. The sum of the areas of the substrate and the product served as an internal standard. The limit of quantification was at least 10 nM (when calculated as the baseline height multiplied by 10).

Novel UPLC method for assaying glutamate carboxypeptidase activity using NAAG as a substrate

The NAAG reaction mixtures were lyophilized for at least 6 h at 20 μbar using Beta 2-8 (LD plus, Christ, Osterode am Harz, Germany) or Labconco catalog number 7753511 (Labconco Instruments Kansas City, MO, USA) and re-dissolved in 25.0 μL of MilliQ water (Merck Millipore, Billerica, MA, USA). Then, surpassing the need for a radioactively labeled substrate traditionally used when assessing the NAAG hydrolysing activity of GCPII [2], a modified fully automated ortho-phthalaldehyde-derivatization [38] step was performed: 11.0 μL of each re-dissolved reaction was manually transferred into a 96-well plate, and the plate was inserted into an autosampler set to 4 °C. A 99.0-μL aliquot of the derivatization solution [40.6 mM 2-mercaptopropionic acid (Sigma-Aldrich), 33.0 mM ortho-phthalaldehyde (Sigma-Aldrich) in 200 mM sodium borate (Pharmacia, Uppsala, Sweden), pH 10.0] was added by the autosampler. After mixing, the sample was injected onto a C18 Acquity UPLC HSS T3 2.1 x 100 mm column with 1.8-μm particles (Waters), guarded by a 0.22-μm pre-filter (Waters) and VanGuard precolumn (Waters), coupled to an Agilent 1200 Series UPLC instrument (Agilent Technologies). Ortho-phthalaldehyde-Glu derivative

was eluted isocratically with 96.0% mobile phase A [25 mM sodium phosphate buffer, pH 6.0 (Penta), supplemented with 0.02% (w/v) sodium azide (Sigma-Aldrich)] and 4.0% mobile phase B (acetonitrile). The column temperature was 70.0 °C. The method consisted of 2.7 min of isocratic flow at 4.0% B, 0.1 min of transition to 80.0% B, 4.7 min at 80.0% B, 0.1 min of transition back to 4.0% B and 8.4 min of re-equilibration. Glutamate eluted after ~ 1.4 min. Analytes were detected by fluorescence at 230/450 nm. The substrate turnover was quantified as the ratio of the peak areas of the glutamic acid (product) and the total cleavage reaction. The limit of quantification was estimated as $x_B + 10\sigma_B$ (mean of the blank measurement plus 10 × SD of the blank) and was usually in the range 20–30 nM.

Thermofluor assay

To determine the temperature midpoint for the protein unfolding transition, T_m, a thermal shift assay using SYPRO® Orange (Invitrogen, Carlsbad, CA, USA) dye [39] was performed on a Roche LightCycler® 480 II instrument (Roche Diagnostics, Penzberg, Germany) in Roche LightCycler® 480 Multiwell Plates 96. Each well contained 3.9 μg of protein and 2.5 μL of 40 x SYPRO® Orange in a final volume of 50 μL of 25 mM Tris (pH 7.5). During pipetting, all components were kept on ice.

Acknowledgements

The use of the Advanced Photon Source was supported by the US Department of Energy (Contract number W-31-109-Eng38) and the use of MX 14.2 was in part funded by the Helmholtz-Zentrum Berlin and BioStruct-X (grant agreement number 283570). C.B. acknowledges support from the EMBO (Installation grant 1978) and IRG (project number 249220). This publication is supported by the project 'BIOCEV – Biotechnology and Biomedicine Centre of the Academy of Sciences and Charles University' (CZ.1.05/1.1.00/02.0109) from the European Regional Development Fund, in part by EU OPPC program CZ.2.16/3.1.00/24016, by the Intramural Research Program of the National Institutes of Health (NIH), National Cancer Institute, Center for Cancer Research, and by grant P304-12-0847 from the Grant Agency of the Czech Republic. The authors would like to thank Pavlina Rezacova for collecting the diffraction data for the rhGCPII-Glu424Ala-FolGlu$_2$ complex at BESSYII (Berlin, Germany), Barbara Slusher (School of Medicine, John Hopkins University, MD, USA) for her kind gift of the 2-PMPA inhibitor, Hillary Hoffman for language help and Radko Soucek for his excellent technical assistance with the amino acid analysis.

Author contributions

MN was responsible for protein cloning, purification and crystallization, structure solving, kinetic and inhibition studies, and preparation of the manuscript. JP was responsible for His475Tyr crystallization and structure solving. PS was responsible for His475Tyr cloning. JS was responsible for protein expression in S2 cells. JL was responsible for X-ray data collection and preparation of the manuscript. CB conceived the project, crystallized proteins and solved the X-ray structures. JK conceived the project, analyzed data and wrote the manuscript.

References

1 Israeli RS, Powell CT, Fair WR & Heston WD (1993) Molecular cloning of a complementary DNA encoding a prostate-specific membrane antigen. *Cancer Res* **53**, 227–230.

2 Robinson MB, Blakely RD, Couto R & Coyle JT (1987) Hydrolysis of the brain dipeptide N-acetyl-L-aspartyl-L-glutamate. Identification and characterization of a novel N-acetylated alpha-linked acidic dipeptidase activity from rat brain. *J Biol Chem* **262**, 14498–14506.

3 Mesters JR, Barinka C, Li WX, Tsukamoto T, Majer P, Slusher BS, Konvalinka J & Hilgenfeld R (2006) Structure of glutamate carboxypeptidase II, a drug target in neuronal damage and prostate cancer. *EMBO J* **25**, 1375–1384.

4 Klusak V, Barinka C, Plechanovova A, Mlcochova P, Konvalinka J, Rulisek L & Lubkowski J (2009) Reaction mechanism of glutamate carboxypeptidase II revealed by mutagenesis, X-ray crystallography, and computational methods. *Biochemistry* **48**, 4126–4138.

5 Barinka C, Rojas C, Slusher B & Pomper M (2012) Glutamate carboxypeptidase II in diagnosis and treatment of neurologic disorders and prostate cancer. *Curr Med Chem* **19**, 856–870.

6 Pavlicek J, Ptacek J & Barinka C (2012) Glutamate carboxypeptidase II: an overview of structural studies and their importance for structure-based drug design and deciphering the reaction mechanism of the enzyme. *Curr Med Chem* **19**, 1300–1309.

7 Pinto JT, Suffoletto BP, Berzin TM, Qiao CH, Lin S, Tong WP, May F, Mukherjee B & Heston WD (1996) Prostate-specific membrane antigen: a novel folate hydrolase in human prostatic carcinoma cells. *Clin Cancer Res* **2**, 1445–1451.

8 Halsted CH, Wong DH, Peerson JM, Warden CH, Refsum H, Smith AD, Nygard OK, Ueland PM, Vollset SE & Tell GS (2007) Relations of glutamate carboxypeptidase II (GCPII) polymorphisms to folate

and homocysteine concentrations and to scores of cognition, anxiety, and depression in a homogeneous Norwegian population: the Hordaland Homocysteine Study. *Am J Clin Nutr* **86**, 514–521.
9 Vargas-Martinez C, Ordovas JM, Wilson PW & Selhub J (2002) The glutamate carboxypeptidase gene II (C>T) polymorphism does not affect folate status in the Framingham Offspring cohort. *J Nutr* **132**, 1176–1179.
10 Chen J, Kyte C, Valcin M, Chan W, Wetmur JG, Selhub J, Hunter DJ & Ma J (2004) Polymorphisms in the one-carbon metabolic pathway, plasma folate levels and colorectal cancer in a prospective study. *Int J Cancer* **110**, 617–620.
11 Devlin AM, Clarke R, Birks J, Evans JG & Halsted CH (2006) Interactions among polymorphisms in folate-metabolizing genes and serum total homocysteine concentrations in a healthy elderly population. *Am J Clin Nutr* **83**, 708–713.
12 Mir MM (2008) Combined impact of polymorphism of folate metabolism genes; glutamate carboxypeptidase, methylene tetrahydrofolate reductase and methionine synthase reductase on breast cancer susceptibility in Kashmiri women. *Clin Chem* **54**, A126.
13 Zhang AX, Murelli RP, Barinka C, Michel J, Cocleaza A, Jorgensen WL, Lubkowski J & Spiegel DA (2010) A remote arene-binding site on prostate specific membrane antigen revealed by antibody-recruiting small molecules. *J Am Chem Soc* **132**, 12711–12716.
14 Dunn B (2012) Proteinases as drug targets., 1st edn. RSC Publishing, Cambridge, UK.
15 Barinka C, Rovenska M, Mlcochova P, Hlouchova K, Plechanovova A, Majer P, Tsukamoto T, Slusher BS, Konvalinka J & Lubkowski J (2007) Structural insight into the pharmacophore pocket of human glutamate carboxypeptidase II. *J Med Chem* **50**, 3267–3273.
16 Devlin AM, Ling EH, Peerson JM, Fernando S, Clarke R, Smith AD & Halsted CH (2000) Glutamate carboxypeptidase II: a polymorphism associated with lower levels of serum folate and hyperhomocysteinemia. *Hum Mol Genet* **9**, 2837–2844.
17 Pokorna J, Machala L, Rezacova P & Konvalinka J (2009) Current and novel inhibitors of HIV protease. *Viruses* **1**, 1209–1239.
18 Barinka C, Mlcochova P, Sacha P, Hilgert I, Majer P, Slusher BS, Horejsi V & Konvalinka J (2004) Amino acids at the N- and C-termini of human glutamate carboxypeptidase II are required for enzymatic activity and proper folding. *Eur J Biochem* **271**, 2782–2790.
19 Luthi-Carter R, Barczak AK, Speno H & Coyle JT (1998) Hydrolysis of the neuropeptide N-acetylaspartylglutamate (NAAG) by cloned human glutamate carboxypeptidase II. *Brain Res* **795**, 341–348.
20 Alquicer G, Sedlak D, Byun Y, Pavlicek J, Stathis M, Rojas C, Slusher B, Pomper MG, Bartunek P & Barinka C (2012) Development of a high-throughput fluorescence polarization assay to identify novel ligands of glutamate carboxypeptidase II. *J Biomol Screen* **17**, 1030–1040.
21 Pietrzik K, Bailey L & Shane B (2010) Folic acid and L-5-methyltetrahydrofolate comparison of clinical pharmacokinetics and pharmacodynamics. *Clin Pharmacokinet* **49**, 535–548.
22 Chen C, Ke JY, Zhou XE, Yi W, Brunzelle JS, Li J, Yong EL, Xu HE & Melcher K (2013) Structural basis for molecular recognition of folic acid by folate receptors. *Nature* **500**, 486–489.
23 Barinka C, Hlouchova K, Rovenska M, Majer P, Dauter M, Hin N, Ko YS, Tsukamoto T, Slusher BS, Konvalinka J *et al.* (2008) Structural basis of interactions between human glutamate carboxypeptidase II and its substrate analogs. *J Mol Biol* **376**, 1438–1450.
24 Hlouchova K, Barinka C, Klusak V, Sacha P, Mlcochova P, Majer P, Rulisek L & Konvalinka J (2007) Biochemical characterization of human glutamate carboxypeptidase III. *J Neurochem* **101**, 682–696.
25 Hlouchova K, Barinka C, Konvalinka J & Lubkowski J (2009) Structural insight into the evolutionary and pharmacologic homology of glutamate carboxypeptidases II and III. *FEBS J* **276**, 4448–4462.
26 Collard F, Vertommen D, Constantinescu S, Buts L & Van Schaftingen E (2011) Molecular identification of beta-citrylglutamate hydrolase as glutamate carboxypeptidase 3. *J Biol Chem* **286**, 38220–38230.
27 Barinka C, Rinnova M, Sacha P, Rojas C, Majer P, Slusher BS & Konvalinka J (2002) Substrate specificity, inhibition and enzymological analysis of recombinant human glutamate carboxypeptidase II. *J Neurochem* **80**, 477–487.
28 Tykvart J, Sacha P, Barinka C, Knedlik T, Starkova J, Lubkowski J & Konvalinka J (2012) Efficient and versatile one-step affinity purification of in vivo biotinylated proteins: expression, characterization and structure analysis of recombinant human glutamate carboxypeptidase II. *Protein Expr Purif* **82**, 106–115.
29 Kibbe WA (2007) OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res* **35**, W43–W46.
30 Barinka C, Starkova J, Konvalinka J & Lubkowski J (2007) A high-resolution structure of ligand-free human glutamate carboxypeptidase II. *Acta Crystallogr F* **63**, 150–153.
31 Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Meth Enzymol* **276**, 307–326.
32 Krug M, Weiss MS, Heinemann U & Mueller U (2012) XDSAPP: a graphical user interface for the convenient processing of diffraction data using XDS. *J Appl Crystallogr* **45**, 568–572.

33 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D* **67**, 235–242.
34 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D* **66**, 486–501.
35 Schuttelkopf AW & van Aalten DM (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* **60**, 1355–1363.
36 Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12–21.
37 Leatherbarrow RJ (1989–2004) GraFit, Version 5.0.11, Erithacus Software Ltd, Horley, UK.
38 Roth M (1971) Fluorescence reaction for amino acids. *Anal Chem* **43**, 880.
39 Ericsson UB, Hallberg BM, DeTitta GT, Dekker N & Nordlund P (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* **357**, 289–298.