Biochemical and Biophysical Research Communications 396 (2010) 157–163

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Understanding specificity in metabolic pathways—Structural biology of human nucleotide metabolism

Martin Welin [a], Pär Nordlund [a,b,*]

[a] Structural Genomics Consortium, Karolinska Institutet, 17177 Stockholm, Sweden
[b] Division of Biophysics, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 17177 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 6 April 2010

Keywords:
Structural biology
Metabolic pathway
Nucleotide metabolism
Specificity determinants
Substrate specificity
Negative selectivity

ABSTRACT

Interactions are the foundation of life at the molecular level. In the plethora of activities in the cell, the evolution of enzyme specificity requires the balancing of appropriate substrate affinity with a negative selection, in order to minimize interactions with other potential substrates in the cell. To understand the structural basis for enzyme specificity, the comparison of structural and biochemical data between enzymes within pathways using similar substrates and effectors is valuable.

Nucleotide metabolism is one of the largest metabolic pathways in the human cell and is of outstanding therapeutic importance since it activates and catabolises nucleoside based anti-proliferative drugs and serves as a direct target for anti-proliferative drugs. In recent years the structural coverage of the enzymes involved in human nucleotide metabolism has been dramatically improved and is approaching completion. An important factor has been the contribution from the Structural Genomics Consortium (SGC) at Karolinska Institutet, which recently has solved 33 novel structures of enzymes and enzyme domains in human nucleotide metabolism pathways and homologs thereof. In this review we will discuss some of the principles for substrate specificity of enzymes in human nucleotide metabolism illustrated by a selected set of enzyme families where a detailed understanding of the structural determinants for specificity is now emerging.

1. Introduction

Many of the physical principles of interactions between enzymes and substrates are well established. However, in the complexity of the cell, the evolution of substrate specificity is not just an optimization process of the affinity between a protein and a substrate but also a process of negative selection, in order to minimize interactions with other molecules in the cell. Therefore, to understand the structural basis for substrate specificity of enzymes in the cell, direct comparison of interactions of enzyme with multiple substrates, as well as between enzymes that bind similar substrates should be valuable. These comparisons are preferably done within pathways and enzyme families. The mapping of the structural basis for specificity within pathways and enzyme families is also of great importance to understand drug selectivity. Systematic structural biology of larger metabolic pathways is still challenging but structural genomics projects have recently demonstrated that structural and biochemical work on large numbers of enzymes, even of human origin, is feasible [1] and that such efforts can complement work in traditional structural biology groups to yield fundamental insights into enzyme structure and mechanism.

Nucleotide metabolism is one of the largest metabolic pathways in human cells and progress towards understanding the structures and mechanisms of the enzymes of this pathway has been significant in recent years. Many of the substrates as well as regulators of these enzymes have related chemical structures and the pathway therefore constitutes an interesting object for an integrative analysis of the structural principles for the evolution of substrate specificity.

Nucleotides provide the building blocks for DNA and RNA and are key players in a wide range of cellular signalling and energy transduction events. Cellular nucleotides can be provided from *de novo* synthetic pathways or be supplied by salvage pathways where nucleobases and nucleosides/deoxynucleosides are recycled from nutrients or from degraded DNA and RNA. To provide appropriate levels of cellular nucleotides, a subset of the enzymes of the human nucleotide metabolism are regulated by allosteric feedback from metabolites of the pathway. At least 14 different conditions of inborn errors in human nucleotide metabolism have been identified where the distortions of the nucleotide pools are the direct cause for the disease [2]. Anti-proliferative nucleoside based drugs constitute one of the most important single classes of therapeutics and human nucleotide metabolism activates and catabolises

* Corresponding author at: Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-17177 Stockholm, Sweden. Fax: +46 8 52486850.
E-mail address: Par.Nordlund@ki.se (P. Nordlund).

0006-291X/$ - see front matter © 2010 Published by Elsevier Inc.
doi:10.1016/j.bbrc.2010.04.054

nucleoside based drugs, as well as being a direct target for several such drugs. As an example more than 60% of all anti-viral drugs are nucleoside based and several important anti-metabolite drugs for cancer and autoimmune diseases are based on the nucleoside scaffold. Nevertheless, many of these drugs are associated with problematic side effects and a more detailed understanding of the mechanism for substrate specificity, as well as for the allosteric regulation, of relevant enzymes, could potentially add momentum to ongoing efforts to improve nucleoside based drugs [3]. Alternatively, structure based drug design targeting key enzymes in human nucleotide metabolism could assist to develop non-nucleoside based anti-proliferate compounds which might provide higher specificity and therefore less side effects.

2. Structural biology of the human nucleotide metabolism

In recent years, the structural coverage of the enzymes in human nucleotide metabolism has been dramatically improved and is approaching completion. Many of the enzymes in nucleotide metabolism from prokaryotic sources are relatively well characterized from a structural and mechanistic viewpoint. However, most of the eukaryotic counterparts have remained less well studied, partially due to the anticipated challenges in pursuing structural studies of these proteins – eukaryotic proteins are generally considered harder to produce and crystallize. Extensive structural information has however been available for some time on a subset of the human enzymes which are particularly important as drug targets or key proteins in the processing of nucleoside based drugs (e.g. IMP-dehydrogenase, thymidylate synthase, nucleoside/nucleotide kinases, phosphorylases and nucleotidases). Since a few years back, the Structural Genomics Consortium (SGC) at Karolinska Institutet has made a significant contribution towards mapping the structural basis for the function of enzymes in human nucleotide metabolism. In this project we are applying a systematic approach to characterize the structure and mechanism of enzymes in human nucleotide metabolism. We have so far deposited 26 structures of different enzymes or enzyme domains belonging to the core human nucleotide metabolism, and additionally seven structures of homologs to these enzymes (http://sgc.ki.se/structures.html). Most of these proteins have been studied in complexes with substrates, substrate analogues, inhibitors or effectors. This effort is benefiting from the establishment of an efficient platform for protein production and crystallization based on the multi-construct strategy [4]. This platform allows many proteins to be studied in parallel and so far ~45 proteins, of the human nucleotide metabolism, or homologs of these proteins, have been targeted by SGC-Stockholm. Success rate improving strategies such as *in situ* proteolysis [5] as well as ligand identification and protein stabilization strategies using thermal shift assays [6,7] have been helpful. To complement the structural studies, systematic biochemical mapping of interactions of these enzymes with compound libraries of physiological nucleotides and nucleotide metabolism intermediates, as well as nucleoside based drugs, is being pursued.

Human nucleotide metabolism can be divided into the *de novo* pathways for purine and pyrimidine synthesis on the one hand, and the salvage and catabolic pathways on the other. To exemplify progress in nucleotide metabolism, an overview of the progress in mapping the structures of the *de novo* pathways is shown in Fig. 1. The *de novo* synthesis of purines where IMP is made from PRPP and small metabolites engages 10 enzyme moieties. Among these enzymes are one tri-functional (GART) and two bi-functional (PAICS and ATIC) and three mono-functional (PPAT, PFAS, and ADSL) enzymes [8]. Except for PPAT and PFAS, all the human structures of the individual enzyme domains in this pathway have been determined. In the subsequent steps toward the production of RNA and DNA precursors, IMP is converted into either GMP or AMP. Four enzymes are responsible for this conversion; inosine-5′-monophosphate dehydrogenase 1 (IMPDH1) and GMP synthase (GMPS) for GMP synthesis and adenylosuccinate synthetase isozyme 2 (ADSS) and adenylosuccinate lyase (ADSL) for AMP synthesis. Now all of these four enzymes have been structurally characterized. In spite of the progress in mapping the structures of the individual components of the purine metabolism, several of these proteins are multi enzyme proteins. Furthermore, it was recently proposed that a number of the purine *de novo* synthesis enzymes are assembled into a larger complex, the "purinosome" [9]. Therefore, the structural characterization of the higher order assemblies of these proteins as well as the implication of these assemblies on the regulation and activity of these enzymes remains to be determined.

The initial steps in the *de novo* pyrimidine pathway are carried out by the CAD protein. CAD is a large tri-functional enzyme composed of three individual domains, glutamine-dependent carbamoyl-phosphate synthase, aspartate carbamoyltransferase and dihydroorotase. The final step in making orotate is catalyzed by dihydroorotate dehydrogenase (DHODH) [10]. So far structural data regarding the human CAD protein are not available, whereas the structure of DHODH has been determined [11]. The subsequent conversion of orotate to UMP is catalyzed by UMP synthase (UMPS), which is a bi-functional enzyme composed of an orotate phosphoribosyltransferase domain and an orotidine 5′-phosphate decarboxylase domain. For DNA synthesis UMP initially has to be phosphorylated to UTP, which is the substrate for CTP synthase (CTPS) converting UTP to CTP. Recent structural efforts have yielded structures of both UMPS domains. Furthermore, the synthase domains of CTPS and its homolog CTPS2 have been determined. Several of these enzymes are up-regulated in tumour cells and thus constitute potential drug targets [12].

Although a number of enzyme components of the purine and pyrimidine pathways have been previously determined, it is clear that the SGC effort has been efficient in complementing these efforts towards a full structural coverage of the *de novo* pathways.

In the salvage pathways nucleobases and ribonucleosides/deoxyribonucleosides (N/dN) are recycled from nutrients and degraded DNA and RNA. Reversible transformations with the nucleobase pools are made both at the nucleoside and nucloside mono-phosphate (NMP) level, by phosphorylases and phosphoribosyltransferases respectively. The catabolic pathways for pyrimidines start from uracil while the catabolic pathways for purines originate from hypoxanthine. Also for the core salvage pathways many of the human enzymes are now structurally characterized.

Due to the size of the nucleotide metabolism and the striking progress in mapping the structural basis for the action of these enzymes only a small fraction of the determined structures can be discussed in this mini-review. To illustrate how the nucleotide substrate can be recognized by the enzymes of the human nucleotide metabolism, we will give examples from four representative enzyme families emphasizing one aspect of specificity in each family.

3. Nucleoside kinases—purine vs. pyrimidine specificity

Nucleoside based drugs need to be activated by phosphorylation and similarly to the N/dN, the phosphorylation of nucleoside analogues assists in trapping them in the cell [13]. There are three nucleoside kinases primarily responsible for phosphorylating N to NMP; Adenosine Kinase (ADK), Uridine–Cytidine Kinase 1 (UCK1) and Uridine–Cytidine Kinase 2 (UCK2). ADK primarily phosphorylates adenosine and deoxyadenosine (dA), while UCK1 and 2 are primarily responsible for the phosphorylation of uridine and cyti-

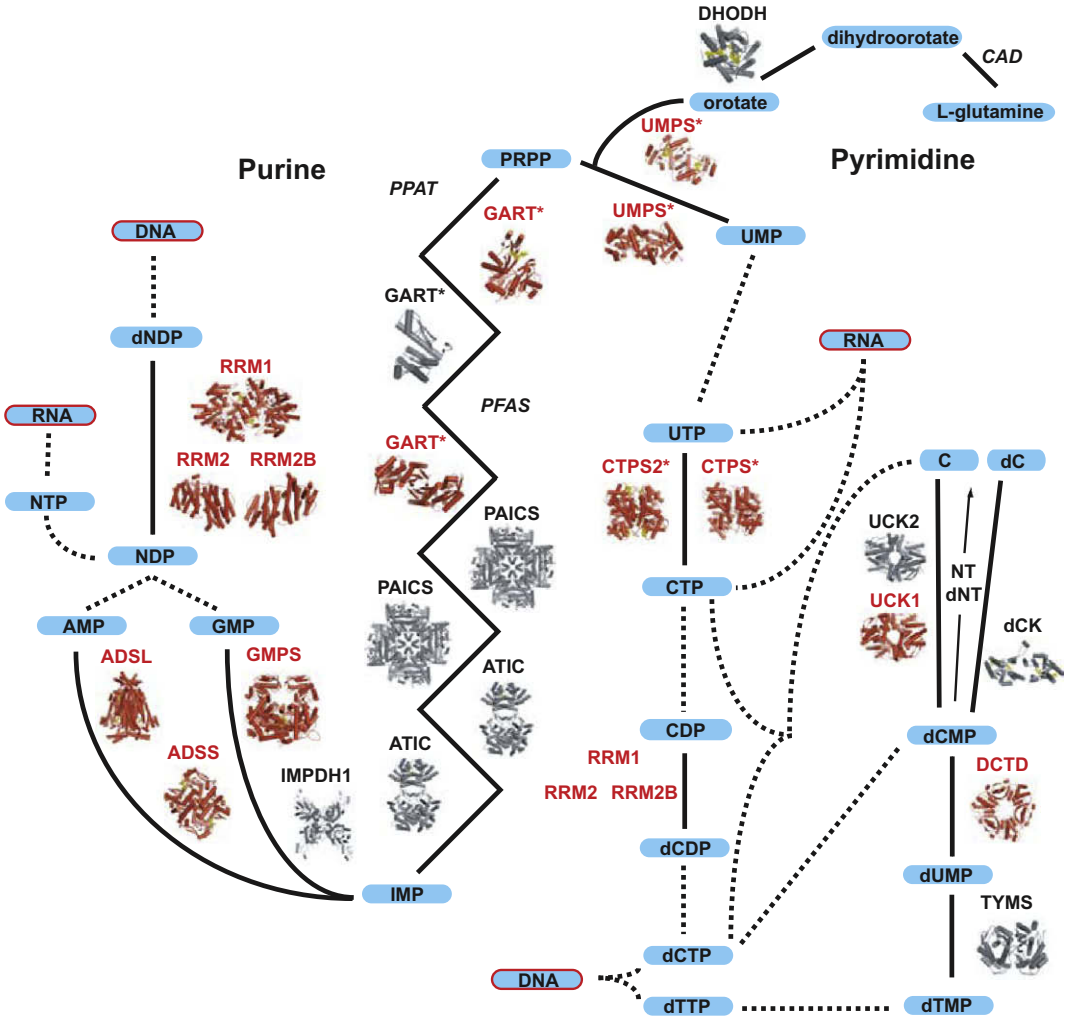

Fig. 1. Schematic view of part of the nucleotide metabolism. Gene names are used for all enzymes. Dotted lines represent enzyme/enzymes that are not specified in the figure. Enzymes coloured in red are structures that have been determined at the Structural Genomics Consortium-Stockholm site. Structures coloured in grey are determined by others and gene names in italic have no human crystal structure. * Corresponds to a structure of one enzymatic step in a multi domain protein. NT and dNT are abbreviations for ribonucleotidase and deoxyribonucleotidase. Enzymes displayed in the purine metabolism, PPAT; Amidophosphoribosyltransferase, GART; Tri-functional purine biosynthetic protein adenosine-3 (pdb-id: 2QK4, 1ZLX, 2V9Y), PFAS; Phosphoribosylformylglycinamidine synthase, PAICS; Multifunctional protein ADE2 (pdb-id: 2H31), ATIC; Bi-functional purine biosynthesis protein PURH (pdb-id: 1PKX), IMPDH1; Inosine-5′-monophosphate dehydrogenase 1 (pdb-id: 1JCN), GMPS; GMP synthase (pdb-id: 2VXO), ADSS; Adenylosuccinate synthetase (pdb-id: 2V40), RRM2; Ribonucleoside-diphosphate reductase subunit M2 (pdb-id: 2UW2), RRM2B; Ribonucleoside-diphosphate reductase subunit M2 B (pdb-id: 2VUX), RRM1; Ribonucleoside-diphosphate reductase large subunit (pdb-id: 2WGH). Enzymes displayed in the pyrimidine metabolism, CAD; CAD protein, Dihydroorotate dehydrogenase (pdb-id: 2B0M), UMPS; Uridine 5′-mono-phosphate synthase (pdb-id: 2WNS, 2JGY), CTPS; CTP synthase 1 (pdb-id: 2VO1), CTPS2; CTP synthase 2 (pdb-id: 3IHL), UCK1; Uridine–cytidine kinase 1 (pdb-id: 2JEO), UCK2; Uridine–cytidine kinase 2 (pdb-id: 1UDW), DCK; Deoxycytidine kinase (pdb-id: 1P60), DCTD; Deoxycytidylate deaminase (pdb-id: 2W4L), TYMS; Thymidylate synthase (pdb-id: 1HZW).

dine to their corresponding NMP [13]. The phosphorylation of dN to dNMP is performed by four different kinases in mammalian cells; deoxycytidine kinase (dCK) and thymidine kinase 1 (TK1), which are located in the cytosol, and deoxyguanosine kinase (dGK) and thymidine kinase 2 (TK2) located in the mitochondria. These enzymes have complementary substrate specificities and are together able to phosphorylate all of the natural deoxynucleosides [14]. The understanding of the structural basis for activity and specificity of the nucleoside-kinases is now relatively well advanced and crystal structures of human dCK, TK1, dGK, ADK, UCK1 (pdb-id: 2JEO), and UCK2 has been determined [15–19]. Insights into their specificity have been obtained by studies of complexes with nucleosides as well as nucleotides with higher phosphorylation states which can serve as feedback inhibitors. The structural basis for specificity has been best described for dN-kinases which will be discussed in the following.

The dN-kinases belong to two structural families. dCK, dGK, and TK2 are part of the same family and have related sequences. Also, the crystal structures of dCK and dGK reveal dimeric proteins with an α/β-domain similar to thymidylate kinase [17]. So far there is no crystal structure of TK2. The second family, to which TK1 belongs, displays a tetramer composed of monomers with two distinct domains; an α/β-domain similar to ATP-binding domains in enzymes of the RecA-F_1ATPase family and a lasso domain stabilized by a structural zinc [16]. TK1 is responsible for phosphorylating thymidine (dT), deoxyuridine (dU) and some nucleoside analogues with modifications at the 5-position of the base and at the 3′-position of the deoxyribose, but have a higher specificity than most other nucleoside-kinases [14]. This high specificity can be explained by the narrow substrate binding pocket of TK1 and that the interactions with the thymine base are coming from main chain atoms, probably making the pocket relatively rigid (Fig. 2C). Together this would not allow sufficient size and

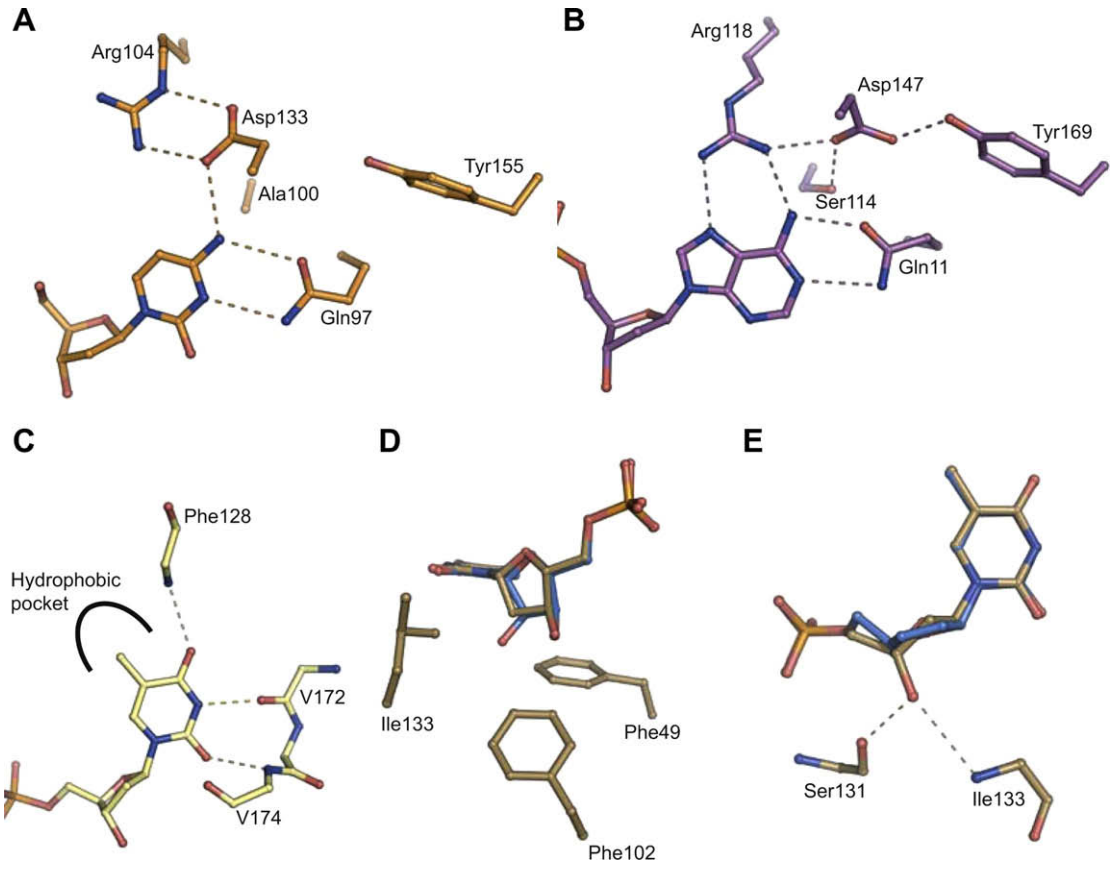

Fig. 2. (A) dCK with bound deoxycytidine (pdb-id: 1P60). Side chains and deoxycytidine are shown with sticks in orange and hydrogen bonds are displayed with orange dotted lines. (B) dGK with bound dATP (pdb-id: 2OCP). Side chains and dATP are shown in magenta and hydrogen bonds are displayed with dotted lines coloured in same colour. (C) TK1 with bound dTTP (pdb-id: 1XBT). Main chain atoms and the thymidine moiety are shown with sticks in yellow and hydrogen bonds are displayed with yellow dotted lines. (D) Comparison of dUMP and UMP binding in dNT2 (pdb-id: 1Z4I, 1Z4M). Side chains are shown with sticks coloured in brown. Substrates dUMP and UMP are shown in brown and blue respectively. (E) Comparison of dTMP and D4TMP binding in dNT2 (pdb-id: 1Z4L, 1Z4Q). Same colouring scheme as in (D) are used.

flexibility for the larger purine nucleoside to bind. Furthermore, the methyl group at the 5-position of thymine base is pointing into a rather small hydrophobic pocket, consistent with that only small substitutions are allowed at this position.

dCK on the other hand has a broad substrate specificity and is able to phosphorylate dC, as well as dA and dG, but the later two with lower efficiency [14]. dCK is also responsible for the activation of several important nucleoside analogues [13] including gemcitabine and cytarabine. The mitochondrial dGK has a higher specificity towards purines and is able to phosphorylate dG and dA, while the activity for pyrimidines is low. The overall structures of dGK and dCK are very similar (root mean square deviation of 1.26 Å), consistent with a relatively high sequence identity of 45%. The structural basis for the differences in substrate specificity can be addressed by comparing the two active sites. Residues involved in binding the base are similar in dCK (Gln97, Ala100, Arg104, Asp133) and dGK (Gln111, Ser114, Arg118, Asp147), where the only difference is the Ala $\Longleftrightarrow$ Ser substitution. The crystal structure of dCK revealed an active site pocket that can accommodate both pyrimidines and purines, explaining a broader specificity of dCK [15]. Discrimination against dT in dCK is explained by the potential steric hindrance between the methyl group at the 5-position and Arg104. Furthermore, an unfavourable interaction between the 4-keto oxygen and Asp133 would be made [15]. In dCK Asp133 is hydrogen bonded to Arg104 and to the amino group of the cytosine base (Fig. 2A). However, in dGK the hydrogen bonding network is slightly different. Ser114 instead of the Ala100 in dCK is hydrogen bonded to Asp147 and thus stabilizes Arg118 in a conformation allowing hydrogen bonding to N6 and N7 of the adenine ring (Fig. 2B). In addition, Asp147 is stabilized in this conformation by interactions with Tyr169. In order for dCK to bind dG, a similar interaction between Asp133 and Tyr155 could be formed. In summary, the structural basis for substrate specificity of the nucleoside-kinases is now relatively well understood, where the size, rigidity and hydrogen bonding networks of the base binding pocket define the substrate preferences.

4. Nucleotidases—ribose specificity

Nucleotidases catalyze the dephosphorylation of mono-phosphate nucleotides producing the corresponding nucleosides [20]. When nucleosides are more readily transported over membranes than the phosphorylated nucleotides, the nucleotidases play a pivotal role in controlling transport events of nucleosides to the exterior of cells. The nucleosides produced by the nucleotidases can also be decomposed to nucleobases and phospho ribose by phosphorylases, which is an entry point to the catabolic pathways for nucleotides. These catabolic processes are also responsible for the degradation of nucleoside based drugs, and the induction of nucleotidase activity can yield cells resistant to such drugs [20]. When the nucleotidases counteracts the activity of the nucleoside-kinases discussed above, the two families form together a so called substrate cycle to control the ratios between (d)N and (d)NMPs. In humans there are six known intracellular soluble 5′-nucleotidases, with varying specificity for different (d)NMPs. Three of these

are regulated by allosteric effectors, although the exact role of this regulation in controlling nucleotide/nucleoside pools remains to be determined. An understanding of the structural basis for the activity and specificity of the human 5′-nucelotidases is now emerging and we have determined the structures of four human soluble 5′-nucleotidases. These are the structure of mitochondrial mdN [21], and more recently the structures of three cytosolic enzymes, cdN, cN-II, and cN-III [22,23]. cN-II is the only one of these enzymes which is allosterically regulated, and the structure allowed one allosteric site to be defined, although the exact mechanism for regulation still remains to be described. MdN have been studied in complex with a battery of physiological substrates as well as nucleotide analogues, which has shed light on the specificity determinants of this family. To allow the trapping of the enzyme-substrate complexes, an inactivating mutant was used where the catalytic aspartate residue (Asp 41 in mdN) was substituted by an asparagine residue. An advantage of the crystallographic methodology is that the binding of relatively low affinity compounds can be studied, i.e. in some instances also the structure of complexes with poor, non-cognate, substrates can be obtained. In mdN the recognition of the nucleotide bases is done in a pocket between two sub-domains of the enzymes [24]. MdN preferably recognise the deoxyribose form of 5′-mono-phosphate nucleotides and the substrate complex structures give some interesting clues on how this ribose specificity is accomplished. In the deoxy form of the 5′-mono-phosphate (dUMP), the 2′-carbon is found in a hydrophobic patch formed by Phe 49, Phe 102, and Ile 133 (Fig. 2D). Interestingly, the ribose containing nucleotide (UMP) binds in a very similar mode to the deoxy form, resulting in exposure of the 2′-OH group of the ribose to the hydrophobic patch formed by Phe 49, Phe 102, and Ile 133. These interactions are unfavourable, which is likely to explain why the ribose form of the nucleotide is a poor substrate. The nucleotide analogue D4TMP is a worse substrate than dTMP, although it differs only in the ribose structure, lacking the 3'-OH group (Fig. 2E). A comparison of the binding mode of the two nucleotides reveals that for dTMP the 3′-OH group makes nice hydrogen bonds to Ser 131 and the main chain amide of Ile 133. D4TMP binds in a mode very similar to dTMP, but lacks the favourable 3′-OH interactions. Instead the 3′-carbon is exposed to this polar environment which is energetically unfavorable, which explains why D4TMP is a poor substrate. In summary, the structural basis for the ribose specificity of the nucleotidases appears to depend on the conserved binding mode of the ribose moiety of different nucleotides to the enzymes, imposing steric hindrance or unfavourable interactions in the ribose binding pocket.

5. Inosine triphosphatase—stringent substrate specificity

Inosine nucleotide triphosphates (ITP or dITP) are present in cells at low levels as by-products generated either by oxidative amination of purine bases or by phosphorylation of IMP. They are removed by the enzyme inosine triphosphate pyrophosphohydrolase (ITPA) which catalyzes the conversion of inosine triphosphates to inosine mono-phosphate and pyrophosphate: (d) ITP → (d) IMP + PP_i. ITPA also scavenges XTP. The primary role for ITPA activity is to prevent the cytosolic accumulation of ITP and dITP which may be incorporated into RNA and DNA. Consistent with this, ITPA has been shown to be important for maintaining chromosome stability [25].

Human ITPA is a homodimeric protein which is reliant upon Mg^{2+} for catalytic activity. While ITP is the best substrate, dITP and XTP are converted with similar kinetics. However, triphosphate nucleotide precursors for DNA and RNA are not substrates for the enzyme [26]. The specificity of ITPA is in fact unusually stringent as compared to other enzymes of nucleotide metabolism, which probably is critical to minimize undesired degradation of the pools of other nucleotides in the cell. Genetic studies of individuals with detectable ITP levels in red blood cells linked ITPA deficiency to several single nucleotide polymorphisms in the ITPA gene. Such mutations have been related to adverse drug reactions of thiopurine based drugs, although the relevance of this association is still being debated [27]. Recently it was shown that patients with ITPA deficiency being given the standard treatment for Hepatitic C virus, including the nucleoside analogue ribavirin, were drastically less prone to develop haemolytic anaemia, a condition often associated with this treatment [28]. Although, the exact mechanism for this effect remains to be determined this suggests that ITPA could be a potential drug target to minimize the effect of such treatments.

We have determined the structure of the human ITPase alone and in complex with one of its physiological substrate ITP [29]. The structures reveal that the binding of the substrate induces a closure of the nucleotide binding grove and leads to a tight interaction of the enzyme with the substrate base which is likely to be important for the high specificity. The closure of the nucleotide binding cleft also positions a number of key side chains for catalysis in their productive conformations, thereby activating the enzyme. Such substrate based activation of the catalytic site is a common mechanism to minimize background activities of enzymes and thereby add additional specificity.

In the structure with bound ITP the purine base is wedged between the hydrophobic side chains of Phe149 and Trp151 (Fig. 3A). The keto oxygen in position 6 of the inosine ring makes extensive interactions with the enzyme where three residues are within hydrogen bonding distance of this oxygen; Lys172, His177, and Arg178. These interactions are probably important to precisely position the nucleotide base in the pocket. From the structure it is clear that an amino group in the position of the keto oxygen would interfere with the hydrogen bonding network, explaining the selectivity of the enzyme against adenine nucleotides. Xanthine binding may be explained by a putative hydrogen bond between the 3-keto oxygen of the xanthine base and the Asp152 backbone amide (not shown). This pocket, however, would not be suitable to accommodate an amino group, explaining the selectivity against a guanine base. An intriguing property of the base interaction is the accumulation of charged residues around the 6-keto oxygen, where at least two of the three residues interacting with the 6-keto group are likely to carry a positive charge. One speculative possibility is that this interaction induces and stabilizes a resonance form of the purine base, where a negative charge develops on the 6-keto group of ITP, which could provide an additional explanation of the high specificity. In summary, the tight interactions with the base in the closed form of the enzymes and the strong interaction with the 6-keto group, together with an unfavourable environment for an amino group at the 3′-position could explain the unusually high specificity of ITPA.

6. Ribonucleotide reductase—allosteric specificity regulation

A key enzyme in nucleotide metabolism is ribonucleotide reductase (RNR) which catalyzes the *de novo* formation of deoxyribonucleotides for DNA synthesis and repair by reducing the ribonucleotide 2′-carbon–oxygen bond [30]. The enzyme is unusual in two ways; it utilizes radical chemistry to catalyze this challenging reaction, and its allosteric regulation involves the control of substrate specificity. The human enzyme belongs to the Class I RNRs which are composed of two dimeric subunits: the R1 subunit which carries the active site and the allosteric effector binding sites, and an R2 subunit which generates and harbours a tyrosine

radical required for the reaction. The R2 protein is induced in S-phase and combines with R1 to provide the dNTPs required for cell division. An alternative R2 protein, P53R2, can be induced out of S-phase to combine with R1 to provide dNTPs for mitochondrial replication and potentially for DNA repair [31]. The human R1 subunit has two types of allosteric sites; an overall activity site where ATP serves as a positive effector and dATP as a negative effector, and a substrate specificity site where the specificity of the enzyme is regulated by feedback from the dNTP pools. The resulting balance of the dNTP pools produced is critical for fidelity of the cellular polymerases and therefore to minimize mutations during replication. The specificity regulation of RNR is unique and constitutes a wonderful example of nature's molecular ingenuity. The regulatory scheme for specificity control can somewhat simplistically be described as a sequential process where the enzyme without effector, or with bound dATP, preferably reduces the pyrimidine substrates UDP and CDP, resulting in elevated levels of dTTP. Binding of dTTP to the enzyme changes the substrate preference to GDP, and in a final step elevated dGTP levels promote ADP reduction.

The initial characterization of the structural basis for specificity regulation of RNRs was done on prokaryotic proteins leading to the identification of the binding site of the specificity effectors and loop 2 as a potential key player in regulating the specificity [32]. A detailed mechanism for the specificity regulation was first revealed for the *Thermotoga martima* enzyme [33] and more recently in the yeast enzyme [34], where structures of cognate pairs of effectors and substrate were determined. The mechanism of the regulation is based on the effector induced conformational modulation of loop 2, which promotes the preferential binding of NDPs. Although the principle of the regulation is similar in the two well characterized enzymes, the structural fine tuning of this regulation is achieved differently, with different loop 2 conformations induced. We have recently solved structures of an N-terminal truncated form of the human R1 subunit with dATP or dGTP bound in the specificity site. The overall structure of the human enzyme is very similar to the yeast enzyme. The conformational modulation of loop 2 due to binding of dATP and dGTP in the human enzyme is quite distinct (Fig. 3B) and is again different to the modulations seen in the yeast enzyme. A common feature of structures of RNRs from different species is that the phosphate groups of NDP (or NTP) are bound in a rigid pocket (not shown). This fixes the substrate and might therefore be important to allow loop 2 to induce the specific recognition by direct interactions with the substrate base. However, the detailed understanding of the structural basis for this modulation in the human RNR has to await further structural studies of cognate pairs of effectors and substrates.

7. Conclusion

Although the energetics for the productive binding of substrates is considerably more complex than indicated in the present review, some patterns are emerging when comparing substrate binding in the human nucleotide metabolism. The nucleobases are often wedged into pockets in-between hydrophobic patches. In some cases, the formation of such pockets is induced by substrate binding, as in the case of ITPA discussed above, where also the activation of the catalytic machinery is concomitant with substrate binding. The specific polar interactions made by the enzymes with the non-solvent exposed parts of the nucleobases do provide hydrogen bonding to all polar groups of the base, although they are sometimes mediated by water molecules. For less specific enzymes, hydrogen bonding between the protein and the bases allows for significant flexibility of coordination, as seen from complex studies with several different substrates (e.g. for dC-kinase discussed above). Many of the substrates of nucleotide metabolism contain phosphate groups, and the structural motifs interacting with phosphate groups. Interactions with the phosphate group(s) of the substrates is an important mean for defining the positioning of the substrate in the active site, although phosphate binding often involves conformational rearrangements due to the involvement of weakly bound divalent ions, most often Mg^{2+}, flexible side chains of arginine and lysine residues, or so called P-loops containing multiple glycine residues in phosphate binding. Another observation not specifically discussed above is that interaction with ribose groups is surprisingly often made by carboxylate side chains of Asp or Glu, forming hydrogen bonding to one or both of the ribose OH-groups. A key process in the evolution of substrate specificity is the negative selection of unwanted physiological substrates. The emerging view from the characterization of enzymes of human nucleotide metabolism is that for the more specific enzymes, such selection is often done by binding poorer substrates

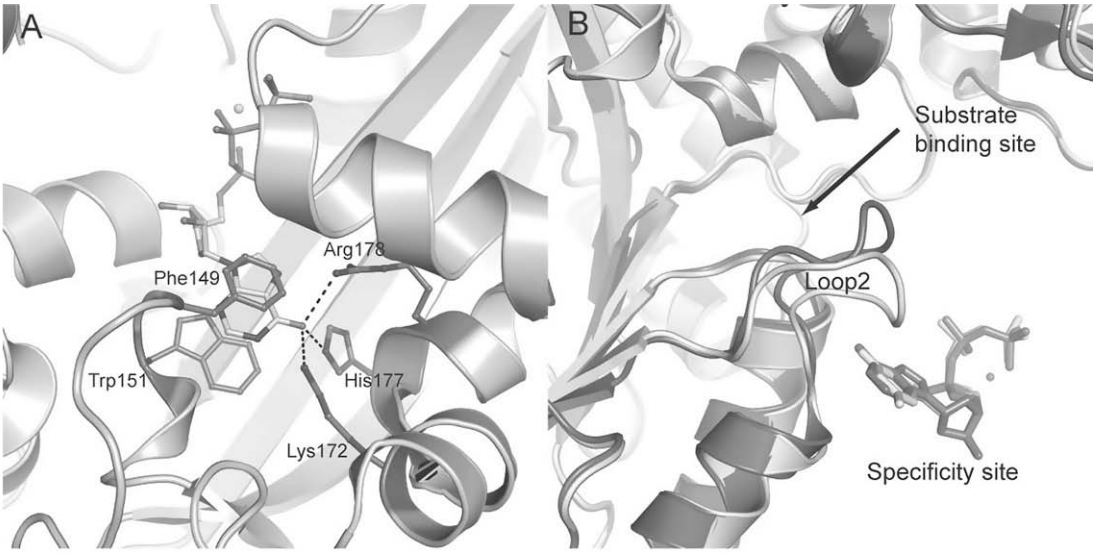

Fig. 3. (A) ITP binding in ITPA (pdb-id: 2J4E). Side chains and ITP are shown with sticks coloured in white and hydrogen bonds are shown with black dotted lines. (B) Cartoon model of comparison of RRM1 in complex with dATP and dGTP bound to the specificity site (pdb-id: 2WGH and unpublished data). RRM1–dATP are coloured in black and RRM1–dGTP in white. The substrate binding site is shown with black arrow. All structure figures were made using Pymol [35].

with similar binding modes as preferred substrates, but at the same time introducing steric or electrostatic hindrance at specific sites to exclude the poorer non-cognate substrates. An integrated view of the structural basis for substrate recognition by the enzymes in human nucleotide metabolism appears to provide some insights for how selectivity for different substrates is achieved, insights which could be useful for future efforts towards developing improved nucleoside based drugs.

Acknowledgments

This work was supported by grants from the Swedish Research Council and the Swedish Cancer Foundation. The Structural Genomics Consortium is a registered charity (No. 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research and Wellcome Trust.

References

[1] A. Edwards, Large-scale structural biology of the human proteome, Annu. Rev. Biochem. 78 (2009) 541–568.
[2] W.L. Nyhan, Disorders of purine and pyrimidine metabolism, Mol. Genet. Metab. 86 (2005) 25–33.
[3] S. Faderl, V. Gandhi, H.M. Kantarjian, Potential role of novel nucleoside analogs in the treatment of acute myeloid leukemia, Curr. Opin. Hematol. 15 (2008) 101–107.
[4] S. Graslund, J. Sagemark, H. Berglund, L.G. Dahlgren, A. Flores, M. Hammarstrom, I. Johansson, T. Kotenyova, M. Nilsson, P. Nordlund, J. Weigelt, The use of systematic N- and C-terminal deletions to promote production and structural studies of recombinant proteins, Protein Expr. Purif. 58 (2008) 210–221.
[5] A. Dong, X. Xu, A.M. Edwards, C. Chang, M. Chruszcz, M. Cuff, M. Cymborowski, R. Di Leo, O. Egorova, E. Evdokimova, E. Filippova, J. Gu, J. Guthrie, A. Ignatchenko, A. Joachimiak, N. Klostermann, Y. Kim, Y. Korniyenko, W. Minor, Q. Que, A. Savchenko, T. Skarina, K. Tan, A. Yakunin, A. Yee, V. Yim, R. Zhang, H. Zheng, M. Akutsu, C. Arrowsmith, G.V. Avvakumov, A. Bochkarev, L.G. Dahlgren, S. Dhe-Paganon, S. Dimov, L. Dombrovski, P. Finerty Jr., S. Flodin, A. Flores, S. Graslund, M. Hammerstrom, M.D. Herman, B.S. Hong, R. Hui, I. Johansson, Y. Liu, M. Nilsson, L. Nedyalkova, P. Nordlund, T. Nyman, J. Min, H. Ouyang, H.W. Park, C. Qi, W. Rabeh, L. Shen, Y. Shen, D. Sukumard, W. Tempel, Y. Tong, L. Tresagues, M. Vedadi, J.R. Walker, J. Weigelt, M. Welin, H. Wu, T. Xiao, H. Zeng, H. Zhu, In situ proteolysis for protein crystallization and structure determination, Nat. Methods 4 (2007) 1019–1021.
[6] U.B. Ericsson, B.M. Hallberg, G.T. Detitta, N. Dekker, P. Nordlund, Thermofluor-based high-throughput stability optimization of proteins for structural studies, Anal. Biochem. 357 (2006) 289–298.
[7] M. Vedadi, F.H. Niesen, A. Allali-Hassani, O.Y. Fedorov, P.J. Finerty Jr., G.A. Wasney, R. Yeung, C. Arrowsmith, L.J. Ball, H. Berglund, R. Hui, B.D. Marsden, P. Nordlund, M. Sundstrom, J. Weigelt, A.M. Edwards, Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination, Proc. Natl. Acad. Sci. USA 103 (2006) 15835–15840.
[8] Y. Zhang, M. Morar, S.E. Ealick, Structural biology of the purine biosynthetic pathway, Cell. Mol. Life Sci. 65 (2008) 3699–3724.
[9] S. An, R. Kumar, E.D. Sheets, S.J. Benkovic, Reversible compartmentalization of de novo purine biosynthetic complexes in living cells, Science 320 (2008) 103–106.
[10] D.R. Evans, H.I. Guy, Mammalian pyrimidine biosynthesis: fresh insights into an ancient pathway, J. Biol. Chem. 279 (2004) 33035–33038.
[11] S. Liu, E.A. Neidhardt, T.H. Grossman, T. Ocain, J. Clardy, Structures of human dihydroorotate dehydrogenase in complex with antiproliferative agents, Structure 8 (2000) 25–33.
[12] G. Weber, Ordered biochemical program of gene expression in cancer cells, Biochemistry (Moscow) 66 (2001) 1164–1173.
[13] A.R. Van Rompay, M. Johansson, A. Karlsson, Substrate specificity and phosphorylation of antiviral and anticancer nucleoside analogues by human deoxyribonucleoside kinases and ribonucleoside kinases, Pharmacol. Ther. 100 (2003) 119–139.
[14] S. Eriksson, B. Munch-Petersen, K. Johansson, H. Eklund, Structure and function of cellular deoxyribonucleoside kinases, Cell. Mol. Life Sci. 59 (2002) 1327–1346.
[15] E. Sabini, S. Ort, C. Monnerjahn, M. Konrad, A. Lavie, Structure of human dCK suggests strategies to improve anticancer and antiviral therapy, Nat. Struct. Biol. 10 (2003) 513–519.
[16] M. Welin, U. Kosinska, N.E. Mikkelsen, C. Carnrot, C. Zhu, L. Wang, S. Eriksson, B. Munch-Petersen, H. Eklund, Structures of thymidine kinase 1 of human and mycoplasmic origin, Proc. Natl. Acad. Sci. USA 101 (2004) 17970–17975.
[17] K. Johansson, S. Ramaswamy, C. Ljungcrantz, W. Knecht, J. Piskur, B. Munch-Petersen, S. Eriksson, H. Eklund, Structural basis for substrate specificities of cellular deoxyribonucleoside kinases, Nat. Struct. Biol. 8 (2001) 616–620.
[18] I.I. Mathews, M.D. Erion, S.E. Ealick, Structure of human adenosine kinase at 1.5 A resolution, Biochemistry 37 (1998) 15607–15620.
[19] N.N. Suzuki, K. Koizumi, M. Fukushima, A. Matsuda, F. Inagaki, Structural basis for the specificity, catalysis, and regulation of human uridine–cytidine kinase, Structure 12 (2004) 751–764.
[20] S.A. Hunsucker, B.S. Mitchell, J. Spychala, The 5′-nucleotidases as regulators of nucleotide and drug metabolism, Pharmacol. Ther. 107 (2005) 1–30.
[21] A. Rinaldo-Matthis, C. Rampazzo, P. Reichard, V. Bianchi, P. Nordlund, Crystal structure of a human mitochondrial deoxyribonucleotidase, Nat. Struct. Biol. 9 (2002) 779–787.
[22] K. Wallden, A. Rinaldo-Matthis, B. Ruzzenente, C. Rampazzo, V. Bianchi, P. Nordlund, Crystal structures of human and murine deoxyribonucleotidases: insights into recognition of substrates and nucleotide analogues, Biochemistry 46 (2007) 13809–13818.
[23] K. Wallden, P. Stenmark, T. Nyman, S. Flodin, S. Graslund, P. Loppnau, V. Bianchi, P. Nordlund, Crystal structure of human cytosolic 5′-nucleotidase II: insights into allosteric regulation and substrate recognition, J. Biol. Chem. 282 (2007) 17828–17836.
[24] K. Wallden, B. Ruzzenente, A. Rinaldo-Matthis, V. Bianchi, P. Nordlund, Structural basis for substrate specificity of the human mitochondrial deoxyribonucleotidase, Structure 13 (2005) 1081–1088.
[25] N. Abolhassani, T. Iyama, D. Tsuchimoto, K. Sakumi, M. Ohno, M. Behmanesh, Y. Nakabeppu, NUDT16 and ITPA play a dual protective role in maintaining chromosome stability and cell growth by eliminating dIDP/IDP and dITP/ITP from nucleotide pools in mammals, Nucleic Acids Res (2010), doi:10.1093/nar/gkp1250.
[26] S. Lin, A.G. McLennan, K. Ying, Z. Wang, S. Gu, H. Jin, C. Wu, W. Liu, Y. Yuan, R. Tang, Y. Xie, Y. Mao, Cloning, expression, and characterization of a human inosine triphosphate pyrophosphatase encoded by the itpa gene, J. Biol. Chem. 276 (2001) 18695–18701.
[27] J. Bierau, M. Lindhout, J.A. Bakker, Pharmacogenetic significance of inosine triphosphatase, Pharmacogenomics 8 (2007) 1221–1228.
[28] J. Fellay, A.J. Thompson, D. Ge, C.E. Gumbs, T.J. Urban, K.V. Shianna, L.D. Little, P. Qiu, A.H. Bertelsen, M. Watson, A. Warner, A.J. Muir, C. Brass, J. Albrecht, M. Sulkowski, J.G. McHutchison, D.B. Goldstein, ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C, Nature 464 (2010) 405–408.
[29] P. Stenmark, P. Kursula, S. Flodin, S. Graslund, R. Landry, P. Nordlund, H. Schuler, Crystal structure of human inosine triphosphatase. Substrate binding and implication of the inosine triphosphatase deficiency mutation P32T, J. Biol. Chem. 282 (2007) 3182–3187.
[30] P. Nordlund, P. Reichard, Ribonucleotide reductases, Annu. Rev. Biochem. 75 (2006) 681–706.
[31] G. Pontarin, A. Fijolek, P. Pizzo, P. Ferraro, C. Rampazzo, T. Pozzan, L. Thelander, P.A. Reichard, V. Bianchi, Ribonucleotide reduction is a cytosolic process in mammalian cells independently of DNA damage, Proc. Natl. Acad. Sci. USA 105 (2008) 17801–17806.
[32] U. Uhlin, H. Eklund, Structure of ribonucleotide reductase protein R1, Nature 370 (1994) 533–539.
[33] K.M. Larsson, A. Jordan, R. Eliasson, P. Reichard, D.T. Logan, P. Nordlund, Structural mechanism of allosteric substrate specificity regulation in a ribonucleotide reductase, Nat. Struct. Mol. Biol. 11 (2004) 1142–1149.
[34] H. Xu, C. Faber, T. Uchiki, J.W. Fairman, J. Racca, C. Dealwis, Structures of eukaryotic ribonucleotide reductase I provide insights into dNTP regulation, Proc. Natl. Acad. Sci. USA 103 (2006) 4022–4027.
[35] W.L. DeLano, The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, 2002.