Research Accomplishments
Dinshaw J. Patel Laboratory

This document summarizes research accomplishments from the Dinshaw J. Patel lab at the Memorial Sloan-Kettering Cancer Center over the last decade (2003-2014) in the fields of (1) riboswitches and ribozymes, (2) RNA silencing pathways, (3) histone mark-mediated epigenetic regulation, (4) DNA methylation mark-mediated epigenetic regulation, (5) pattern recognition receptors of cytosolic nucleic acids, (6) non-vesicular trafficking by lipid transfer proteins, (7) protein-RNA complexes on disease-related systems and (8) DNA lesion architecture and processing of damage sites by bypass polymerases. Color images of the structures can be found by visiting the web site address of the Patel lab: http://www.mskcc.org/mskcc/html/10829.cfm.

Last updated: Dec 15, 2014

1. Riboswitches and Ribozymes

We highlight below our key contributions to structural biology research on the structure, dynamics and catalytic mechanisms of riboswitches and ribozymes. Research on sensing domains of riboswitches with bound metabolites offer the opportunity to elucidate principles underlying higher order folding and stability of compact RNA architectures, as well as how RNA composed of only four nucleotides can form pockets that accommodate specific ligands and discriminate against closely-related analogs. Most sensing domains of riboswitches are composed of helical stems converging on internal loop junctions, with the latter zippering up through maximization of base stacking and hydrogen bonding on complex formation with bound ligands. Small ligands such as planar purines are encapsulated through maximization of hydrogen bond formation in a single three-helical junctional pocket within the overall tuning-fork like architecture, with bound guanine distinguished from adenine through Watson-Crick pairing with a single complementary pyrimidine of the riboswitch (Serganov et al. Chem. Biol. 2004). Larger ligands such as thiamine pyrophosphate (TPP) use an alternate principle by binding their riboswitches in an extended conformation with each end of the metabolite anchored in separate pockets on adjacent stems within the three-helical tuning-fork architecture. Notably, the electrostatic repulsion between the negatively-charged phosphates of the metabolite and riboswitch RNA was overcome by a pair of bridging hydrated divalent Mg^{2+} cations that coordinate the diphosphate of TPP and base edges of the riboswitch, but not the phosphate backbone of the RNA (Serganov et al. Nature 2006). The flavin mononucleotide (FMN) six-helical junctional riboswitch, that adopted a butterfly-like fold, also capitalized on one Mg^{2+} to mediate coordination between the monophosphate of FMN and base edges of the RNA (Serganov et al. Nature 2009).

The five-helical junctional lysine riboswitch instead uses a monovalent K^{+} cation to anchor a hydrogen bond network encapsulating the charged amide end of the bound amino acid (Serganov et al. Nature 2008). The elongated junctional binding pocket for lysine within the riboswitch can accommodate the bound lysine side chain anchored at both its ends in an extended alignment, thereby capitalizing on a distance ruler to discriminate against all other amino acids. The glycine riboswitch is unique given the requirement for generating a binding pocket that can accommodate this smallest of amino acids and the involvement of tandem sensing domains that bind glycine in a cooperative manner. Our structural studies have highlighted the role of a pair of bound hydrated Mg^{2+} cations in extending shape complementarity by filling cavities within the glycine binding pocket and have identified quaternary interactions mediated by stacked A-minor triples between adjacent sensing domains, thereby accounting for the observed binding cooperativity (Huang et al. Mol. Cell 2010).

We have solved the structure of c-di-AMP bound to the T. tengcongensis ydaO riboswitch, thereby identifying a five-helical scaffold containing a zipper-up bubble, a pseudoknot and long-range tertiary base pairs (Ren and Patel, Nat. Chem. Biol. 2014). Highlights include the identification of two c-di-AMP binding pockets on the same face of the riboswitch, related by pseudo two-fold symmetry, with potential for cross-talk between sites mediated by adjacent base stacking alignments connecting pockets. The adenine rings of bound c-di-AMP molecules are wedged between bases and stabilized by stacking, base-sugar and sugar-sugar intermolecular hydrogen bonding interactions.
Finally, our studies establish that the fluoride riboswitch adopts a pseudoknot RNA scaffold to encapsulate this small negatively-charged anion and discriminate against other halides. The fluoride anion is coordinated by three Mg$^{2+}$ cations, which in turn are coordinated by five inwardly-pointing phosphate oxygens (Ren et al. *Nature* 2012). The fluoride riboswitch achieves formation of these concentric shells of opposing charge and unprecedented inward directionality of backbone phosphates by maximizing stacking but not base pairing within the zippered up junctional segment.

Our knowledge of RNA catalysis by natural ribozymes has been limited to the making and breaking of phosphodiester bonds. To this end, we have solved the structure of the env22 twister ribozyme, which adopts a compact tertiary fold stabilized by co-helical stacking, double-pseudoknot formation and long-range pairing interactions (Ren et al. *Nat. Commun.* 2014; with the Ronald Micura lab at the University of Innsbruck, Austria). The U-A cleavage site adopts a splayed-apart conformation with the modeled 2'-O of U positioned for in-line attack on the adjacent to-be-cleaved P-O5' bond. Both an invariant guanosine and a Mg$^{2+}$ are directly coordinated to the non-bridging phosphate oxygens at the U-A cleavage step, with the former positioned to contribute to catalysis and the latter to structural integrity.

We have also elucidated the structure of an in vitro selected 49-mer ribozyme that catalyzes carbon-carbon bond formation between the diene anthracene and the dienophile maleimide, both in the free state and when bound to one stereoisomer of the bicyclic product (Serganov et al. *Nat. Struct. Mol. Biol.* 2005; with Andres Jaschke lab, Heidelberg University). The ribozyme adopts a $\lambda$-shaped nested double pseudoknot fold with a preformed junctional hydrophobic pocket that accommodates the chiral bicyclic Diels-Alder product. Stereoselection is dictated by shape complementarity, augmented by a combination of intermolecular stacking and a limited number of hydrogen bond interactions, with Mg$^{2+}$ cations playing a structural but not catalytic role.


2. RNA Silencing Pathways

Our structural studies of the RNA silencing pathway have focused on the RNase III enzyme Dicer that cleaves long double-stranded RNA (dsRNA) into small interfering RNAs (siRNAs), the RNase H enzyme Argonaute (Ago) that capitalizes on a bound guide strand generated from processed siRNA to cleave complementary messenger RNA (mRNA) targets, and on viral suppressors that target and inhibit distinct steps in the RNA silencing pathway. The existing paradigm in the field is that Dicer uses its PAZ domain to anchor the 3'-end of the bound dsRNA and initiate cleavage to generate siRNAs of defined length based on the fixed separation between the PAZ domain and the composite pocket generated by a pair of interacting RNase III domains. Our unanticipated discovery of a basic 5'-phosphate-binding pocket within the platform domain of Dicer (Tian et al. *Mol. Cell* 2014) provided support for biochemical experiments that established a 5'-counting rule where the distance ruler involved not just 3'-end recognition (3'-counting rule) but more importantly recognition of the 5'-phosphate of bound siRNAs and miRNAs (Park et al. *Nature* 2011; championed by the Narry Kim lab, Seoul National University). Our structural studies of *K. polysporus* Dicer from budding yeast yielded results that overturned another paradigm wherein all Dicers were predicted to use a mechanism that involved processing initiation from the ends and working inwards into the dsRNA. By contrast, our structure-function studies establish that *Kp* Dicer forms dimers that bind cooperatively along the dsRNA substrate, such that the distance between consecutive active sites determines the length of the siRNA products, thereby initiating processing within the interior and working outwards into the dsRNA (Weinberg et al. *Cell* 2011; with David Bartel lab, Whitehead and MIT).

Our pioneering studies of eubacterial Argonautes bound to 5'-phosphorylated DNA guide (Yuan et al. *Mol Cell* 2009; with Thomas Tuschl lab, Rockefeller) and complementary RNA targets of various lengths have provided unprecedented insights into individual steps of the Ago-mediated silencing pathway. Structural studies of the binary complex of *T. thermophilus* Argonaute (*Tt*Ago) bound to 5'-phosphorylated guide DNA identified the nucleic acid-binding channel spanning all domains within Ago, with the 5-phosphorylated and 3'-ends anchored in Mid (Ma et al. *Nature* 2005) and PAZ (Ma et al. *Nature* 2004) domain pockets respectively, and the Watson-Crick edges of stacked bases 2 to 6 spanning the seed segment directed outwards and available for nucleation with complementary bases of the RNA target strand (Wang et al. *Nature* 2008a). The guide strand remained anchored at both its ends for ternary complexes containing RNA targets strands of 12-nt in length, but the 3'-end of the target strand was released for RNA targets of 15-nt in length due to topological constraints associated with formation of longer A-form duplexes within the interior of the Ago scaffold (Wang et al. *Nature* 2008b). The transition from a cleavage incompatible to a cleavage compatible conformation required not just the release of the 3'-end, but also accompanying conformational changes in three loops within Ago that are associated with positioning of a pair of Mg$^{2+}$ cations between the three catalytic aspartate residues from the PIWI domain and the cleavable phosphate on the RNA target strand (Wang et al. *Nature* 2009; with Thomas Tuschl lab). Finally, functional experiments demonstrated that *Tt*Ago acts as a barrier for the uptake and propagation of foreign DNA, thereby functioning in host defense by a DNA-guided DNA interference pathway (Swarts et al. *Nature* 2014a; championed by the John van der Oost lab, Wageningen University, The Netherlands). Our studies have highlighted in unprecedented detail the nucleation, propagation and cleavage steps of Ago-mediated cleavage of target RNAs, representing a striking triumph of structural biology in addressing mechanistic issues related to the RNA silencing pathway.

Our structure-function studies on the complex of eukaryotic budding yeast *K. polysporus* Ago with bound endogenous 5'-phosphorylated guide RNA identified an additional requirement for generation of a catalytically-competent pocket. The transition from an inactive to active conformation involved formation of a hydrogen-bonded network that stabilized an expanded and repositioned loop, which inserts an invariant glutamate into the catalytic pocket, thereby completing a universally conserved RNase H-type catalytic tetrad (Nakanishi et al. *Nature* 2012; with David Bartel lab). Of the four human Agos, only hAgo2 exhibits slicer activity. Based on a comparison of binary complexes of hAgo1 and hAgo2 bound to endogenous guide RNAs, selective substitutions of a subset of amino acids in hAgo1 generated a measureable level of slicer activity (Nakanishi et al. *Cell Reports* 2014; with Thomas Tuschl lab).
We have reviewed structure-function and phylogenetic studies and analysis of prokaryotic and eukaryotic Agos, thereby allowing a reconstruction of the evolutionary journey of the Argonaute proteins through the three domains of life and how they relate to their distinct physiological roles (Swarts et al. Nat. Struct. Mol. Biol. 2014b; with John van der Oost lab and Eugene Koonin lab, National Library of Medicine, NIH)

Functional studies championed in the group of Alexei Aravin lab (Caltech) have concluded that transcriptionally inherited piRNAs act as an epigenetic memory for identification of substrates for piRNA biogenesis on two levels: by inducing a permissive chromatin environment for piRNA precursor synthesis and by enhancing processing of these precursors (Le Thomas et al Genes Dev. 2014). Specifically, Rhino recruits the piRNA biogenesis factor Cutoff to piRNA clusters and is required for efficient transcription of piRNA precursors.

In a functional study championed by the Mien-Chie Hung lab (M. D. Anderson Cancer Center), the epidermal growth factor receptor (EGFR), which is the product of a well-characterized oncogene in human cancers, suppresses the maturation of specific tumor suppressor-like mRNAs in response to hypoxic stress through phosphorylation of hAgo2 at Tyr393 (Shen et al. Nature 2013). The association between EGFR and hAgo2 is enhanced by hypoxia, leading to elevated hAgo2 Tyr393 phosphorylation, which in turn reduced the binding of Dicer to Ago2 and inhibited miRNA processing from precursor miRNAs to mature miRNAs.

Our structure of the complex between the viral suppressor p19 and siRNA has defined the molecular basis for sequestration of the siRNA by this viral suppressor. Reading-head α-helices projecting from opposite ends of the core architecture formed by the symmetrical p19 homodimer precisely positioned pairs of tryptophan rings for stacking over the terminal base pairs of the bound siRNA, thereby using a caliper-like mechanism to measure the length of the bound siRNA duplex (Ye et al. Nature, 2003). In another study, biochemical approaches have highlighted how the viral suppressor 2b blocks A. thaliana Ago1 cleavage activity to inhibit miRNA pathways, attenuate RNA silencing and counter host defense (Zhang et al. Genes Dev. 2006; championed by the Nam-Hai Chua lab, Rockefeller).

Structural efforts are underway in our lab to complement functional studies championed by the V. Narry Kim lab (Seoul National University) to understand the role of uridylation by TUTases marking mRNA for degradation (Lim et al. Cell, 2014) and adenylation of maternally inherited miRNAs by Wispy (Lee et al. Mol. Cell, 2014).


3. Histone Mark-mediated Epigenetic Regulation

We itemize and highlight below our major accomplishments in the field of histone mark-mediated epigenetic regulation. Our initial efforts have focused on deciphering molecular principles underlying readout by writers, readers and erasers of histone marks. We were instrumental in pioneering studies that identified PHD fingers (Li et al. Nature 2006; with David Allis lab, Rockefeller) and BAH domains (Kuo et al. Nature 2012; with Or Gozani lab, Stanford) as readers of lysine methylation marks using an aromatic cage capture mechanism. Follow up studies by our group identified three distinct modes of methyl-lysine (Kme) recognition, namely ‘surface groove’ (Li et al. Mol. Cell 2007; with David Allis lab) and ‘cavity insertion’ (Li et al. Mol. Cell 2007; with David Allis lab) modes defined by cation-π interactions within aromatic-lined pockets and the ‘composite junctional’ mode (Iwase et al. Nat. Struct. Mol. Biol. 2011; with Yang Shi lab, Harvard Medical School) formed at the interface of two adjacently positioned
reader modules with recognition defined by surface complementarity and atypical hydrogen bond formation. Next, dysregulation of Kme readout was shown to impact on leukemogenesis from studies on the PHD finger of JARID1A fused to nucleoporin-98 (Wang et al. *Nature* 2009; with David Allis lab) and on primordial dwarfism from studies on the BAH domain of origin of replication complex protein-1 (Kuo et al. *Nature* 2012; with Or Gozani lab).

Structure-function studies of simultaneous readout of Kme and acetyl-lysine (Kac) marks by the PHD-Bromo cassettes of TRIM24 (Tsai et al. *Nature* 2010; with Michelle Barton lab, M. D. Anderson Cancer Center) and TRIM33 (Xi et al. *Cell*, 2011; with Joan Massague lab, Memorial Sloan-Kettering Cancer Center) highlighted how both binding affinity and specificity are enhanced during multivalent readout of dual marks at the histone tail level. Related more challenging studies on the BPTF PHD-Bromo cassette established principles underlying multivalent readout at the nucleosomal level (Ruthenburg et al. *Cell* 2011; with David Allis lab). Independently, the PHD-Bromo cassette of MLL1 was identified as a regulatory platform orchestrating cyclophilin Cyp33-mediated cis-trans proline isomerization, in a process facilitating histone deacetylase recruitment, thereby switching MLL1 from an activated to a repressive state (Wang et al. *Cell* 2010; with David Allis lab).

KDM2A, a member of the jumonji C domain-containing histone lysine demethylase family, has been shown to specifically targets lower methylation states (me1/me2) of H3K36. Our structural studies reveal that H3K36-specificity for KDM2A is mediated by the U-shaped threading of the H3K36 peptide through a catalytic groove on KDM2A (Cheng et al. *Genes Dev.* 2014; with Or Gozani lab, Stanford University). Further, structure-guided substitutions of residues in the KDM2A catalytic pocket abrogate KDM2A-mediated functions important for suppression of cancer cell phenotypes. Our efforts to target Kme-binding pockets in erasers capitalized on a structure-guided and chemoproteomics approach that identified a small molecule catalytic site inhibitor of the H3K27me3-specific KDM6 subfamily jumonji lysine demethylases (Kruidenier et al. *Nature* 2012; championed by the GaxoSmithKline lab, Stevenage, UK).


We have also successfully targeted the tetrameric helical alignment of the NHR2 domain of AML1-ETO, a fusion protein generated by translocation in acute myeloid leukemia, through complex formation with the N2B motif of E proteins, thereby defining the fundamental principle whereby transcription factors create a new protein-binding interface through oligomerization (Sun et al. *Nature* 2013; with Robert Roeder lab, Rockefeller).

Our studies on histone chaperones have focused on the histone-binding domain of DAXX, which though unstructured in the free state, adopts a folded conformation composed of six α-helices on complex formation with histone H3.3/H4, thereby encapsulating the histone dimer, and in the process competing with major inter-histone, DNA and other chaperone interactions sites (Elsasser et al. *Nature* 2012; with David Allis lab).

Mammalian chromatin undergoes reorganization demarcated by H2A.X Ser139 phosphorylation (γ-H2A.X) during the double-strand break response. Our functional studies on the Williams-Beuren syndrome transcription factor (WSTF), a component of the WICH chromatin-remodeling complex, has identified a new regulatory mechanism mediated by phosphorylation of Tyr142 of H2A.X (Xiao et al. *Nature* 2009; with the David Allis lab). Phosphorylation was observed at either Ser139 or Tyr142, but not both.

In a separate project, our efforts have focused on identifying proteins that target the p11-AnxA2 complex, as an approach towards contributing mechanistic insights to our current understanding of the signaling pathway governing selective serotonin reuptake inhibitors (SSRI). To this end, our structure-function studies established a central role for SMARCA3, a chromatin-remodeling factor targeted to the p11/annexin A2 heterotetrameric complex, in the SSRI/p11 signaling pathway (Oh et al. *Cell* 2013; championed by the Paul Greengard lab, Rockefeller). These studies suggest an approach to the development of improved antidepressant therapies.


4. DNA Methylation Mark-mediated Epigenetic Regulation

Our structure-function studies on readout of DNA methylation marks have focused on maintenance DNA methyltransferases in mammals (DNMT1) and plants (CMT3). Mammalian DNMTases, like their bacterial counterparts, contain a methyltransferase domain, but also contain additional RFD, CXXC and a pair of BAH domains. Our structural studies on two distinct DNMT1-DNA complexes defined the positioning of these domains relative to the bound DNA, as well as the alignment of the catalytic pocket containing the methyl donor S-adenosylmethionine relative to the to-be-methylated flipped-out cytosine. Our studies have established how a combination of autoinhibitory (Song et al. Science 2011) and productive (Song et al. Science 2012) mechanisms ensures the high fidelity of DNMT1-mediated maintenance DNA methylation.

We have solved the dimeric structure of the methyltransferase domain of DRM, a key de novo methyltransferase in plants (Zhong et al. Cell, 2014; collaboration with the Steve Jacobsen lab, UCLA Medical School). Functional studies establish that Arabidopsis DRM2 exists in complex with siRNA effector Ago4 and preferentially methylates one DNA strand, likely the strand acting as the template for RNA pol-V-mediated non-coding RNA transcripts. These data support a model in which DRM2 is guided to target loci by siRNA-Ago4 and involves base-pairing of associated siRNAs with nascent RNA transcripts.

Functional studies championed by the Steve Jacobsen lab have established that non-CG methylation in Arabidopsis is governed by the methyltransferase CMT2 which preferentially binds H3K9me2 and methylates non-CG cytosines that are regulated by H3K9 methylation (Stroud et al. Nat. Struct. Mol. Biol., 2014). These studies demonstrate extensive dependencies of small RNA accumulation and H3K9 methylation patterning on non-CG methylation, suggesting self-reinforcing mechanisms between these epigenetic factors.

In Arabidopsis, CHG DNA methylation is controlled by the H3K9 methylation mark through a self-reinforcing loop between DNA methyltransferase CMT3 and H3K9 histone methyltransferase Kryptonite (KYP). Structure-function studies on ZMET2, a maize analog of CMT3, which stably associates with
H3K9me2-containing nucleosomes, established that the aromatic cage pockets of its BAH and Chromo domains independently targeted H3K9me2 marks, with a complete loss in binding activity in vivo on mutation of either aromatic-lined pocket (Du et al. Cell 2012; with Steve Jacobsen lab). These structural insights complemented by in vivo functional data on key mutants establish how the H3K9me2 mark on histone tails recruits CMT3 to the DNA substrate for methylation. We have also solved the structure of KYP in complex with methylated DNA, substrate H3 peptide and cofactor SAH, thereby defining the spatial positioning of the SRA domain relative to the SET domain (Du et al. Mol. Cell 2012; with Steve Jacobsen lab). The methylated DNA is bound by the SRA domain with the 5mC flipped out of the DNA, while the H3(1-15) peptide substrate binds between the SET and post-SET domains, with the ε-ammonium of K9 positioned adjacent to bound cofactor analog SAH. These structural insights complemented by in vivo functional data on key mutants of residues lining the 5mC and H3K9-binding pockets within KYP, establish how methylated DNA recruits KYP to the histone substrate for methylation. Together, the structures of KYP and CMT3 complexes provide insights into molecular mechanisms linking DNA and histone methylation.

Establishment of all DNA methylation and maintenance of some non-CG methylation in A. thaliana involves the RNA-directed DNA methylation pathway. There are two main steps in this pathway, which ultimately result in the targeting of the DNA methyltransferase DRM2. One step, traditionally thought to be upstream, involves the synthesis of 24-nt siRNAs via the RNA polymerase pol-IV complex and associated proteins, primarily SSH1. The second, or downstream step, involves the production of non-coding transcript by pol-V, in association with SUVH2/9, whereby they are thought to act as scaffolds that interact with siRNAs to eventually recruit DRM2 and its associated DNA methyltransferase activity. Our structure-function studies established that recruitment of pol-IV to a large subset of sites required SHH1, whose tandem Tudor domains bind to unmodified H3K4 and methylated H3K9, with the latter a silencing mark that is also found at RNA-directed DNA methylation sites (Law et al. Nature 2013, with Steve Jacobsen lab). Our structure-function studies on the linkage between SUVH2/9 and pol-V established that this RNA polymerase is recruited to DNA methylation sites through methyl-DNA binding SUVH2/9 proteins, with our mechanistic findings outlining a means for selectively targeting regions of the plant genome for epigenetic silencing (Johnson et al. Nature 2014; with Steve Jacobsen lab). Thus, the dual marks of DNA methylation and H3K9 methylation appear to act in a self-reinforcing loop mechanism to maintain pol-IV and pol-V activity at sites of RNA-directed DNA methylation.


5. Pattern Recognition Receptors of Cytosolic Nucleic Acids

Much attention has been drawn to the field of pattern recognition receptors of cytosolic nucleic acids with the recent identification of cGAS as a metazoan sensor of dsDNA in the cytosol, whereby it catalytically converts GTP and ATP to cGAMP, with this second messenger in turn targeting the adaptor hSTING, thereby triggering a cascade of events leading to interferon production and an innate immune response. The structural biology efforts of our group successfully identified how cGAS undergoes a conformational change on binding dsDNA to generate a catalytically-competent binding pocket capable of accommodating dinucleotide intermediate pppGpA and product cGAMP. In a paradigm shift in the field, our group were the first to demonstrate that cGAMP contains an unanticipated 2',5'-linkage at the GpA step and a standard 3',5' linkage at the ApG step, thereby identifying a founding member of a family of metazoan 2',5'-containing cyclic dinucleotides distinct from bacterial all 3',5' cyclic dinucleotide counterparts (Gao et al. Cell 2013a; with Thomas Tuschl lab, Winfried Barchet and Gunther Hartmann labs, University Bonn-Hospital and Roger Jones lab, Rutgers).

Extension of this research to complex formation of hSTING with bound linkage isomers of cGAMP established that both binding and cellular activities were maximal for 2',5'-linkage isomers. Given that the antiviral drug DMXAA targets mSTING but not hSTING, our group identified three mutations within hSTING that rendered DMXAA sensitive to hSTING, thereby opening opportunities for structure-guided modification of DMXAA as an effective drug for treatment of viral-related diseases in humans (Gao et al. Cell 2013b; Gao et al. Cell Reports 2014; with Tuschl, Barchet, Liang Deng [MSKCC] and Jones labs).

Our structural studies on the RIG-I C-terminal domain bound to blunt-end double-stranded 5'-ppp RNA established how a lysine-rich basic cleft sequesters the 5'-triphosphorylated end of the bound RNA (Wang et al. Nat. Struct. Mol. Biol. 2010; with Gunther Hartmann and Thomas Tuschl labs).


6. Non-vesicular Trafficking by Lipid Transfer Proteins

We have extended our earlier contributions to the structural biology of glycosphingolipids (GSLs) bound to glycolipid transfer proteins (GLTPs), to that of ceramide-1-phosphate (C1P) bound to newly-identified ceramide-1-phosphate transfer protein (CPTP). The structures of both families of complexes explain how the head-group surface-groove binding pockets distinguish between neutral GSLs and negatively-charged C1P ligands. In addition, structural studies of GSL-GLTP (Malinina et al. Nature 2004; with Rhoderick Brown lab, Hormel Institute) and C1P-CPTP (Simanshu et al. Nature 2013a; with Rhoderick Brown lab and Charles Chalfont lab, Virginia Commonwealth University) complexes identified details of the alignment of one or both lipid chains in a molded-to-fit hydrophobic channel, with partitioning between sphingosine-in and sphingosine-out binding modes dependent on acyl chain length and
occupancy of the hydrophobic tunnel (Malinina et al. *PLoS Biol*. 2006; with Rhoderick Brown lab). The structural data are consistent with a cleft-like conformational gating mechanism, whereby glycolipid chains sequentially enter and leave the tunnel in the membrane-associated state during membrane vesicle biogenesis and trafficking, signal transduction and immunological presentation processes.

We have also demonstrated the structure-based design of ‘engineered GLTPs’ with enhanced specificity for select GSLs (Samygina et al *Structure* 2010; with Lucy Malinina lab, CIC bioGUNE, Bilbao and Rhoderick Brown lab), thereby opening opportunities for new therapeutic approaches towards targeting GSL-mediated pathologies. More recently, we have demonstrated that the *Arabidopsis* Accelerated Cell Death 11 (ACD11) protein is a C1P transfer protein and intermediary reglator of phytoceramide levels (Simanshu et al. *Cell Reports*, 2014; with Rhoderick Brown lab).


7. Protein-RNA Complexes on Disease-related Systems


Of notable interest are the intermolecular pairing alignments between the guanidinium group of arginines and base edges of guanosines, together with surface complementarity, to the specificity and stability of complex formation. These structures have highlighted the diversity of principles underlying molecular recognition between protein and RNA, including the role of protein dimers serving as scaffolds that provide a pair of RNA-binding sites, thereby facilitating a chain-reversal looping mechanism for sequestration of RNA targets containing tandem sequence recognition elements.

Recently, we have extended our studies of protein-RNA complexes to their protein-DNA counterparts. We have identified the Ben domain as a novel sequence-specific DNA-binding domain conserved in neural transcriptional repressors (Dai et al. *Genes Dev.*, 2013; with Eric Lai lab, Memorial...
Sloan-Kettering Cancer Center), and in a follow-up paper, outlined common and distinct DNA-binding and regulatory activities of the Ben-solo transcription factor family (Dai et al. *Genes Dev.*, 2014; championed by the Eric Lai lab).


8. DNA Lesion Architecture and Processing of Damage Sites by Bypass Polymerases

NMR-based structural studies by our group on chiral polycyclic aromatic hydrocarbon (PAH) ‘bay’ region benzo[a]pyrene-N²-guanosine (BP-N²G) and ‘fjord’ region benzo[c]phenanthrene-N⁶-adenosine (BPh-N⁶A) adducts bound covalently to DNA have resulted in the highly original discovery of three distinct structural alignments adopted by these lesions on DNA (with Nicholas Geacintov and Suse Broyle labs, New York University). These are minor groove alignment of BP-N²G lesion without disruption of the modified base pair (Cosman et al. *Proc. Natl. Acad. Scis. USA* 1992), intercalation of BP-N²G lesion with modified base displacement and disruption of modified base pair (Cosman et al. *Biochemistry* 1993a), and intercalation of a propeller-like non-planar BPh-N⁶A lesion without disruption of modified base pair (Cosman et al. *Biochemistry* 1993b). Further, inversion of chirality at all four positions on the benzylic ring of the PAH results in opposingly-oriented alignments of the lesions in all three structural categories (reviewed in Geacintov et al. *Chem. Res. Toxicol.* 1997). The distribution of these three alignment
families was determined by the chiral characteristics of individual stereoisomers within the benzylic ring of the PAHs and by whether the covalent adducts are directed towards the minor or the major groove edges of the DNA helix. These conformational differences for PAH alignments impact on the recognition and rates of excision of these lesions by repair enzyme systems.

In a parallel crystallographic structural research program (with Nicholas Geacintov and Suse Broyde labs), we have addressed the consequences of processing of oxidative, alkylation damage and aromatic amine adducts (reviewed in Patel et al. Chem. Res. Toxicol. 1998) of guanosine in DNA by Dpo4 bypass polymerase. Our efforts have yielded mechanistic insights in unprecedented detail into the translocation mechanics mediated by Dpo4 bypass polymerase during a cycle of binding and incorporation of nucleoside triphosphates opposite the oxoG lesion (Rechkoblit et al. PLoS Biol. 2006), as well as the consequences of error-free and error-prone bypass opposite the aminofluorene-C8-guanosine (AF-C8-G) lesion (Rechkoblit et al. Nat. Struct. Mol. Biol. 2010).


