Research Accomplishments (Dinshaw J. Patel Laboratory)

This document summarizes research accomplishments from the Dinshaw J. Patel lab at the Memorial Sloan-Kettering Cancer Center primarily in the last 14 years (2003-2016) in the fields of (1) RNA architecture, recognition and catalysis, (2) RNA-mediated gene regulation, (3) CRISPR-Cas cleavage of dsDNA, (4) histone mark-mediated epigenetic regulation, (5) histone chaperones, (6) DNA methylation mark-mediated epigenetic regulation, (7) cGAS-cGAMP-STING innate immune response pathway, (8) lipid transfer proteins, (9) protein-RNA complexes on disease-related systems and (10) DNA damage and processing 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.

A profile of Dinshaw J. Patel focusing on his scientific achievements is available at *Proc. Natl. Acad. Scis. USA* 112, 10570-10572 (2015).

Last updated: Dec 29, 2016

1. RNA Architecture Recognition and Catalysis

We highlight below our key contributions to structural biology research on the structure and dynamics of ligand-sensing domains of riboswitches and catalytic mechanisms of self-cleaving ribozymes.

Riboswitches

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 adjacently-aligned stems within the three-helical tuning-fork architecture (Serganov et al. *Nature* 2006). 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. 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 L-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 L-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 quarternary interactions mediated by stacked A-minor triples between adjacently-aligned sensing domains, thereby accounting for the observed binding cooperativity (Huang et al. *Mol. Cell* 2010). Our combined x-ray, NMR and MD studies of the L-glutamine riboswitch have identified a L-glutamine-dependent conformational transition from tuning fork to L-shaped alignment of stem segments (Ren et al. *Cell Reports* 2015a). Our studies identified an open ligand-binding pocket that combined with a high conformational penalty for forming the ligand-bound state, provide mechanisms for reducing L-glutamine binding affinity while retaining high selectivity.

We have studied the structure of the *pfl* riboswitch that selectively binds the alarmone ZMP (5-amino-4-imadazole carboxamide ribose-5'-monophosphate) and regulates genes associated with purine biosynthesis and one-carbon metabolism. The riboswitch adopts a pseudoknot-based scaffold, with molecular recognition involving shape complementarity, supplemented by a network of intermolecular hydrogen bonds and an unanticipated direct coordination of the carboxamide of ZMP by a Mg^{2+} cation (Ren et al. *Structure* 2015c).

We have solved the structure of 3',3'-c-di-AMP bound to the *T. tengcongenesis ydaO* riboswitch, thereby identifying a five-helical scaffold containing a zippered-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 adjacently-aligned 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. These studies on cyclic dinucleotide second messengers have been extended to structural and functional studies of 3',3'-cGAMP bound to its riboswitch in *Geobacter*, whereby the riboswitch adopts a tuning-fork architecture with a junctional ligand-binding pocket to accommodate the bound ligand in a defined orientation (Ren et al *Cell Reports* 2015b; with Ming Hammond laboratory, UC-Berkeley). The specificity of ligand recognition was shown to be affected by point mutations outside the binding pocket

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.

We have written two reviews on the structure and dynamics of ligand-sensing domains of riboswitches (Serganov & Patel, 2007; Serganov & Patel, 2012).

Huang, L., Serganov, A. & Patel, D. J. (2010). Structural insights into ligand recognition by a sensing domain of the cooperative glycine riboswitch. ***Mol. Cell*** 40, 774-786.

Pikovskaya, O., Polonskkaya, A., Patel, D. J. & Serganov, A. (2011). Structural principles of nucleoside selectivity in a 2'-deoxyguanosine. ***Nat. Chem. Biol.*** 7, 748-755.

Ren, A. and Patel, D. J. (2014). c-di-AMP binds the *ydaO* riboswitch in two pseudo-symmetry-related pockets. ***Nat. Chem. Biol*****.** 10, 780-786.

Ren, A., Rajashankar, K. & Patel, D. J. (2012). Fluoride ion encapsulation by Mg^{2+} and phosphates in a fluoride riboswitch. ***Nature*** 486, 85-89.

Ren, A., Rajashankar, K. R. and Patel, D. J. (2015c). Global fold and molecular recognition for the *pfl* riboswitch bound to ZMP, a master regulator of one carbon metabolism. ***Structure*** 23, 1375-1381.

Ren, A., Wang, X. C., Kellenberger, C. A., Rajashankar, K. R., Jones, R., Hammond, M. C. and Patel, D. J. (2015b). Structural basis for molecular discrimination by a 3',3'-cGAMP riboswitch. ***Cell Reports*** 11, 1-12.

Ren, A., Xue, Y., Peselis, A., Serganov, A., Al-Hashimi, H. and Patel, D. J. (2015a). Structural and dynamic basis for low-affinity, high-selectivity binding of L-glutamine by the glutamine riboswitch. ***Cell Reports*** 13, 1800-1813.

Serganov, A. & Patel, D. J. (2007). Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. ***Nat. Rev. Genetics*** 8, 776-790.

Serganov, A. & Patel, D. J. (2012). Metabolite recognition principles and molecular mechanisms underlying riboswitch function. ***Ann. Rev. Biophys*****.** 41, 343-370.

Serganov, A., Huang, L. & Patel, D. J. (2008). Structural insights into amino acid binding and gene control by a lysine riboswitch. ***Nature*** 455, 1263-1267.

Serganov, A., Huang, L. & Patel, D. J. (2009). Coenzyme recognition and gene regulation by a FMN riboswitch. ***Nature*** 458, 233-237.

Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R. & Patel, D. J. (2006). Structural basis for gene regulation by a thiamine pyrophosphate-binding riboswitch. ***Nature*** 441, 1167-1171.

Serganov, A., Yuan, Y-R., Pikovskaya, O., Polonskaia, A., Malinina, L., Phan, A. T., Hobartner, C., Micura, R., Breaker, R. R. & Patel, D. J. (2004). Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. ***Chem. Biol.*** 11, 1729-1741.

Ribozymes

Our knowledge of RNA catalysis by natural ribozymes has been limited to the making and breaking of phosphodiester bonds. To this end, we have 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 λ-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.

Our group has also been interested in elucidating the catalytic cleavage mechanisms adopted by small nucleolytic self-cleaving ribozymes. To this end, we have focused on the twister and pistol ribozymes identified in the Ronald breaker laboratory. 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 either structural integrity and/or catalysis. These studies have been extended in research championed by the Micura laboratory to identify a mini-twister ribozyme variant and to use biochemical and labeling approaches to study the impact of catalytic pocket residues and cations on the phosphodiester backbone cleavage by this ribozyme class (Kosutic et al. *Angew. Chemie Int . Edn*. 2015).

We have recently solved the structure of the pistol ribozyme, thereby defining the pseudoknot fold that facilitates site-specific in-line cleavage (Ren et al. *Nat. Chem. Biol.* 2016; with the Ronald Micura lab at the University of Innsbruck, Austria). Our studies also identified highly conserved guanine and adenine residues involved in general acid-base catalysis, with the latter exhibiting an increase in pKa by one pH unit. These studies on the pistol ribozyme defined how he overall and local topologies dictate the in-line alignment at the G-U cleavage site, with cleavage assays on variants revealing key residues that participate in acid-base catalyzed cleavage xhemistry.

Kosutic, M., Neuner, S., Ren, A., Flur, S., Wunderlich, C., Mayrhofer, E., Vusurovic, N., Seikowski, J., Westhof, E., Hobartner, C., Patel, D. J., Kreitz, C. and Micure, R. (2015). A mini-twister variant and impact of residues/cations on the phosphodiester cleavage chemistry of this ribozyme class. ***Angew. Chemie Int. Edn.*** 54, 15128-15133.

Ren, A., Kosutic, M., Rajashankar, K. R., Frener, M., Santner, T., Westhof, E., Micura, R. and Patel, D. J. (2014). In-line alignment and Mg^{2+} coordination at the cleavage site of the twister ribozyme. ***Nat. Commun.*** 5, e5534.

Ren, A., Vusurovic, N., Gebetsberger, J., Gao, P., Juen M., Kreutz, C., Micura, R. & Patel, D. J. (2016). Pistol ribozyme adopts an embedded pseudoknot fold facilitating site-specific in-line self-cleavage. ***Nat. Chem. Biol.*** 12, 702-708.

Serganov, A., Keiper, S., Malinina, L., Tereschko, V., Skripkin, E., Hobartner, C., Polonskaia, A., Phan, A. T., Wombacher, R., Micura, R., Dauter, Z., Jaschke, A. & Patel, D. J. (2005). Structural basis for Diels-Alder ribozyme catalyzed carbon-carbon bond formation. ***Nature Struct. & Mol. Biol.*** 12, 218-224.

2. RNA-mediated Gene Regulation

Our research in RNA-mediated gene regulation has focused on the role of prokaryotic and eukaryotic Argonaute and Piwi proteins in RNA silencing and Cas proteins in CRISPR-Cas mediated cleavage of dsDNA.

RNA Silencing: Argonautes

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 (Wang et al. 2009; Sheng et al. 2014), 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).

Structure-function studies have established that the alpha motif of Tas C-terminus mediates RITS (RNA-induced transcriptional silencing) *cis* spreading and promotes heterochromatic gene silencing (Li et al. *Mol. Cell* 2009).

Li, H., Motamedi, M., Wang, Z., Patel, D. J. & Moazed, D. (2009), An alpha motif of Tas3 C-terminus mediated RITS cis-spreading and promotes heterochromatin gene silencing. ***Mol Cell*** 34, 155-167.

Ma, J-B., Ye, K. & Patel, D. J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. ***Nature*** 429, 318-322.

Ma, J. B., Yuan, Y. R., Meister, G., Pei, Y., Tuschl, T. & Patel, D. J. (2005). Structural basis for 5'-end-specific recognition of the guide RNA strand by the *A. fugidus* PIWI protein. ***Nature*** 434, 666-670.

Nakanishi, K., Ascano, M., Gogakos, T., Ishibi-Murakami, S., Serganov, A. A., Briskin, D., Morozov, P., Tuschl, T. and Patel, D. J. (2013). Eukaryote-specific insertion elements control human ARGONAUTE slicer activity. ***Cell Reports*** 3, 1893-1900.

Nakanishi, K., Weinberg, D. E., Bartel, D. P. & Patel, D. J. (2012). Structure of yeast Argonaute with guide RNA. ***Nature*** 486, 368-374.

Shen, J., Xia, Y., Khotskaya, Y. B., Huo, L., Nakanishi, K., Lim, S-O., Du, Y., Wang, Y., Chang, W-C., Chen, C-H., Hsu, J. L., Lam, Y. C., James, B. P., Liu, C-G., Liu, X., Patel, D. J. & Hung, M. C. (2013). EGFR modulates miRNA maturation in response to hypoxia through phosphorylation of Ago2. ***Nature*** 497, 383-387.

Sheng, G., Zhao, H., Rao, Y., Wang, J., Swarts, D. C, van der Oost, J., Patel, D. J. and Wang, Y. (2014). Structure-based cleavage mechanism of *T. thermophiles* Argonaute DNA guide strand-mediated DNA target cleavage. ***Proc. Natl. Acad. Scis. USA.*** 111, 652-657.

Swarts, D. C., Jore, M. M., Westra, E. R., Zhu, Y., Janssen, J. H., Wang, Y., Patel, D. J., Berenguer, J., Brouns, S. J. and van der Oost, J. (2014a). DNA-guided DNA interference by prokaryotic Argonaute. ***Nature*** 507, 258-261.

Swarts, D. C., Makarova, K., Wang, Y., Nakanishi, K., Ketting, R. F., Koonin, E. V., Patel, D. J. and van der Oost, J. (2014b). The evolutionary journey of Argonaute proteins. ***Nat. Struct. Mol. Biol.*** 21, 743-753.

Wang, Y., Juranek, S., Li, H., Sheng, G., Tuschl, T. & Patel, D. J. (2008b). Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. ***Nature*** 456, 921-926.

Wang, Y., Juranek, S., Li, H., Sheng, G., Wardle, G. S., Tuschl, T. & Patel, D. J. (2009). Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. ***Nature*** 461, 754-761.

Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. (2008a). Structure of the guide-strand-containing argonaute silencing complex. ***Nature*** 456, 209-213.

Ye, K., Malinina, L. & Patel, D. J. (2003). Recognition of siRNA by a viral suppressor of RNA silencing. ***Nature*** 426, 874-878.

Yuan, Y. R., Ma, J. B., Kuryavyi, V., Pei, Y., Zhadina, M., Meister, G., Chen, H. Y., Dauter, Z., Tuschl, T. & Patel, D. J. (2005). Crystal structure of *Aquifex aeolicus* Argonaute provides unique perspectives into the mechanism of guide strand-mediated mRNA cleavage. ***Mol. Cell*** 19, 405-419.

Zhang, X., Yuan, Y-R., Pei, Y., Tuschl, T., Patel, D. J. & Chua, N-H. (2006). Cucumber mosaic virus-encoded 2b suppressor inhibits *Arabidopsis* AGO1 cleavage activity to counter plant defense. ***Genes Dev.*** 20, 3255-3268.

RNA Silencing: Dicers

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-bindng 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).

Park, J. E., Heo, I., Tian, Y., Simanshu, D. K., Chang, H., Jee, D., Patel, D. J. & Kim, V. N. (2011). Dicer recognizes the 5'-end of RNA for efficient and accurate cleavage. ***Nature*** 475, 201-205.

Tian, Y., Simanshu, D. K., Ma, J. B., Park, J-E, Heo, I., Kim, V. N. & Patel, D. J. (2014). A phosphate-binding pocket within the platform-PAZ cassette of human Dicer. ***Mol. Cell*** 53, 606-616.

Weinberg, D., Nakanishi, K., Patel, D. J. & Bartel, D. P. (2011). The inside-out mechanism of Dicers from budding yeasts. ***Cell*** 146, 262-276.

RNA Silencing: piRNA Pathways

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. Additional functional studies championed in the group of Alexei Aravin lab (Caltech) have established that Aub and Ago3 are recruited to nuage through two mechanisms to form a ping-pong complex assembled by Krimper (Webster et al. *Mol. Cell* 2015). In another project championed by the Alexei Aravin lab (Caltech), the protein Cutoff was shown to suppress RNA polymerase II termination to ensure expression of piRNA precursors (Chen et al. *Mol. Cell* 2016).

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 (Ye et al. *Nature*, 2003). 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. 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).

Chen, Y-C. A., Stuwe, E., Luo, Y., Ninova, M., Thomas, A. L., Rozhavskaya, K., Li, S., Vempati, S., Laver, J. D., Patel, D. J., Smibert, C. A., Lipshitz, H. D., Toth, K. F. and Aravin, A. A. (2016). Cutoff suppresses RNA polymerase II termination to ensure expression of piRNA precursors. ***Mol. Cell*** 63, 97-109.

Le Thomas, A., Stuwe, E., LI, S., Du, J.,,Marinov, G., Rozhkov, N., Chen, A. Y-C., Luo, Y., Sachidanandam, R., Tot, K. F., Patel, D. J. and Aravin, A.A. (2014). Trans-generationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. ***Genes Dev*****.** 28, 1667-1680.

Webster, A., Li, S., Hur, J. K., Wachsmuth, M., Bois, J., Perkins, E. M., Patel, D. J. and Aravin, A. A. (2015). Aub and Ago3 are recruited to nuage through two mechanisms to form a ping-pong complex assembled by Krimper. ***Mol. Cell*** 59, 564-575.

RNA Tailing

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; Kim et al. *EMBO J.* 2015) and adenylation of maternally inherited miRNAs by Wispy (Lee et al. *Mol. Cell*, 2014).

Kim, B., Ha, M., Leoff, L., Chang, H., Simanshu, D. K., Li, S., Patel, D. J., Joo, C. and Kim, V. N. (2015). TUT7 controls the fate of precursor miRNAs by using three different uridylation mechanisms. ***EMBO J.*** 34, 1801-1815.

Lee, M., Choi, Y., Kim, K., Jin, H., Lim, J., Nguyen, T. A., Yang, J., Jeong, M., Giraldez, A. J., Yang, H., Patel, D. J. and Kim, V. N. (2014). Adenylation of maternally inherited microRNAs by Wispy. ***Mol Cell*** 56, 696-707.

Lim, J., Ha, M., Chang, H., Kwon, S. C., Simanshu, D. K., Patel, D. J. and Kim, V. N. (2014). Uridylation by TUT4 and TUT7 marks mRNA for degradation. ***Cell*** 159, 1365-1376.

3. CRISPR-Cas Cleavage of dsDNA

CRISPR-Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated protein) provides adaptive immune protection and helps archaea ad bacteria defend themselves against phage infection.

Cas Complexes

The early structural efforts were focused on Cas9, its binary complex with crRNA/tracrRNA as a binary complex and with added target dsDNA as a ternary complex. Recently, two new Cpf1 and C2c1 Cas proteins have been identified that are distinct from Cas9. Both Cpf1 and C2c1 recognize a distal T-rich PAM (protospacer adjacent motif) and cause staggered dsDNA breaks in contrast to Cas9, which recognizes a proximal G-rich PAM and causes blunt end dsDNA breaks.

We have solved the structure of the ternary complexes of Cpf1 endonuclease with crRNA guide and dsDNA target (Gao et al. *Cell Research* 2016) and compared it with the published structure of the binary complex with crRNA. The seed sequence required for initial DNA recognition is disordered in the binary complex, but becomes ordered on ternary complex formation. The A-T pairs of the PAM duplex segment are recognized from both the minor and major grooves, while the PAM-interacting cleft of Cpf1 undergoes an open-to close conformational change on ternary complex formation.

We have also solved the structures of C2c1 bound to crRNA/tracrRNA as a binary complex and with added dsDNA as a ternary complex, thereby capturing catalytically competent conformations with both target and non-target strands independently positioned in a single RuvC catalytic pocket (Yang et al. *Cell* 2016). crRNA adopts a pre-ordered five-nucleotide A-form seed sequence in the binary complex, with release of an inserted tryptophan, facilitating zippering-up of 20-bp RNA-DNA heteroduplex on ternary complex formation. Structural comparison of C2c1 ternary complexes with their Cas9 and Cpf1 counterparts highlights the diverse mechanisms adopted by these distinct CRISPR-Cas9 systems, thereby broadening and enhancing their applicability as genome editing tools.

Gao, P., Yang, H., Rajanshankar, K. R., Huang, Z. & Patel, D. J. (2016). Type V CRISPR-Cas Cpf1 endonuclease employs a unique mechanism for crRNA-mediated target recognition. ***Cell Research*** 26, 901-913.

Yang, H., Gao, P., Rajashankar, K. R., & Patel, D. J. (2017). PAM-dependent target DNA recognition and cleavage by C2c1 CRISPR-Cas endonuclease. ***Cell*** 167,1814-1828.

4. 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.

Single Mark Readout

We were instrumental in pioneering studies that identified PHD fingers (Li et al. *Nature* 2006; Taverna et al. 2006; with David Allis lab, Rockefeller) and BAH domains (Kuo et al. *Nature* 2012; with Or Gozani lab, Stanford), as well as contributors to Tudor domains (Cai et al *Mol. Cell* 2013; with Greg Wang lab at Univ. of North Carolina) 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. *Nature* 2006; 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, as well as 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. Additional studies reported on H3R2me0 recognition by the WDR5 module of MLL1 (Ruthenburg et al. *Nat. Struct. Mol. Biol.* 2006; with Greg Verdine lab at Harvard) and the PHD finger of UHRF1 (Rajakumara et al. *Mol. Cell* 2011; with Yang Shi lab, Harvard Medical School), as well as H3R2 and H4Kme0 recognition by the BAH-PHD cassette of plant ORC1b (Li et al. *Structure* 2016; with Steve Jacobsen lab at UCLA).

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). We have written reviews outlining the readout of epigenetic posttranslational modifications from a structural perspective (Taverna et al. *Nat. Struct. Mol. Biol.* 2007; Patel and Wang, *Ann. Rev. Biochem***.** 2013)

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.

Cai, L., Rothbart, S. B., Lu, R., Xu, B., Tripathy, A., Chen, W-Y., Zheng, D., Patel, D. J., Allis, C. D., Strahl, B. D., Song, J. & Wang, G. (2013). An H3K36me3-containing Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. ***Mol. Cell*** 49, 571-582.

Cheng, Z., Cheung, P., Kuo, A. J., Yuki, E. T., Wilmot, C. M., Gozani, O. and Patel, D. J. (2014). A molecular threading mechanism underlies jumonji lysine demethylase KDM2A regulation of methylated H3K36. ***Genes Dev*****.** 28, 1758-1771.

Iwase, S., Xiang, B., Ghosh, S., Ren, T., Lewis, P. W., Cochrane, J. C., Allis, C. D., Picketts, D. J., Patel, D. J., Li, H. & Shi, Y. (2011). ATRX links atypical histone methylation recognition mechanisms to human cognitive function. ***Nat. Struct. Mol. Biol.*** 18, 769-776.

Kuo, A. J., Song, J., Cheung, P., Ishibe-Murakami, S., Yamazoe, S., Chen, J., Patel, D. J. & Gozani, O. (2012). ORC1 BAH domain links dimethylation of H4K20 to DNA replication licensing and Meier-Gorlin syndrome. ***Nature*** 484, 115-119.

Li, H., Fischle, W., Wang, W. K., Duncan, E. M., Liang, L., Murakami-Ishibe, S., Allis, C. D. & Patel, D. J. (2007). Structural basis for lower lysine methylation state-specific readout by MBT repeats and an engineered PHD finger module. ***Mol. Cell*** 28, 677-691.

Li, H., Ilin, S., Wang, W. K., Wysocka, J., Allis, C. D. & Patel, D. J. (2006). Molecular basis for site-specific readout of H3 lysine 4 trimethylation by the BPTF PHD finger. ***Nature*** 442, 91-95.

Patel, D. J. and Wang, Z. (2013). A structural perspective of readout of epigenetic posttranslational modifications. ***Ann. Rev. Biochem*****.** 82, 81-118.

Rajakumara, E., Wang, Z., Ma, H., Hu, L., Chen, H., Lin, Y., Guo R., Wu, F., Li, H., Lan, F., Shi, Y., Xu, Y., Patel, D. J. & Shi, Y. (2011). PHD finger recognition of unmodified histone H3R2 links UHRF1 to regulation of euchromatic gene expression. ***Mol. Cell*** 43, 275-284.

Ruthenberg, A. J., Wang, W., Graybosch, D. M., Li, H., Allis, C. D., Patel, D. J. & Verdine, G. L. (2006). Histone H3 lysine 4 methylation state recognition and presentation by the WDR5 module of the MLL1 complex. ***Nature Struct. Mol. Biol*****.** 13, 704-712.

Taverna, S. D., Illin, S., Rogers, R. S., Tanny, J. C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L. P., Chait, B., Patel, D. J., Aitchison, J. D., Tackett, A. J. & Allis, C. D. (2006). Yng1 PHD finger binding to H3

trimethylated at K4 targets promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. ***Mol. Cell*** 24, 785-796.

Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. (2007). How chromatin-binding modules interpret histone modifications: Lessons from professional pocket pickers. ***Nat. Struct. Mol. Biol.*** 14, 1025-1040.

Wang, G. G., Song, J., Wang, Z., Dormann, H. L., Casadio, F., Li, H., Luo, J., Patel, D. J. & Allis, C. D. (2009). Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. ***Nature*** 459, 847-851.

Multivalent Mark Readout

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). We have written a review on multivalent engagement of chromatin modifications by linked binding modules (Ruthenburg et al., *Nat. Rev. Mol. Cell Biol.* 2007).

Ruthenburg, A. J., Li, H., Patel, D. J. & Allis, C. D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. ***Nat. Rev. Mol. Cell Biol.*** 8, 983-994.

Ruthenburg, A., Li, H., Milne, T., Dou, Y., McGinty, R. K., Yuen, M., Muir, T. W., Patel, D. J. & Allis, C. D. (2011). Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. ***Cell*** 145, 692-706.

Tsai, W-W., Wang, Z., Yiu, T. T., Akdemir, K. C., Xia, W., Winter, S., Tsai, C-Y., Shi, X., Schwarzer, D., Plunkett, W., Aronow, B., Gozani, O., Fischle, W., Hung, M. C., Patel, D. J. & Barton, M. C. (2010). TRIM24 links recognition of a non-canonical histone signature to breast cancer. ***Nature*** 468, 927-932.

Wang, Z., Song, J., Milne, T. A., Wang, G. G., Li, H., Allis, C. D. & Patel, D. J. (2010). Pro isomerization in MLL1 PHD3-Bromo cassette connects H3K4me3 readout to CyP33 and HDAC-mediated repression. ***Cell*** 141, 1183-1194.

Xi, Q., Wang, Z., Zaromytidou, A., Zhang, X. H., Chow-Tsang, L-F., Liu, J. X., Kim, H., Monova-Todorova, K., Kaartinen, V., Studer, L., Mark, W., Patel, D. J. & Massague, J. (2012). A poised chromatin platform for Smad access to master regulators. ***Cell*** 147, 1511-1524.

Small Molecule Inhibitors

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). We have written a review on small molecule epigenetic inhibitors targeted to lysine methyltransferases and demethylases (Wang and Patel, *Quart. Rev. Biophys*. 2013).

Kruidenier, L., Chung, C., Cheng, Z., Liddle, J., Bantscheff, M., Bountra, C., Bridges, A., Che, K., Diallo, H., Eberhard, D., Hutchinson, S., Joberty, G., Jones, E., Katso, R., Leveridge, M., Mosley, J., Rowland, P., Ramirez-Molina, C., Schofield, C. J., Sheppard, R., Smith, J. E., Swales, C., Tanner, R., Thomas, P.,

Tumber, A., Drewes, G., Oppermann, U., Patel, D. J., Lee, K., & Wilson, W. (2012). A selective H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. ***Nature*** 488, 404-408.

Sun, X-J., Wang, Z., Wang, L., Jiang, Y., Chen, W-Y., Melnick, A., Patel, D. J., Nimer, S. D. & Roeder, R. G. (2013). A stable transcription factor complex nucleated by dimeric AML1-ETO controls leukemogenesis. ***Nature*** 500, 93-97.

Wang, Z. and Patel, D. J. (2013). Small molecule epigenetic inhibitors targeted to histone lysine methyltransferases and demethylases. ***Quart. Rev. Biophys*.** 46, 349-373.

Proteins Regulating Epigenetic Events

A collaborative structure-function study has established that the PZP (PHD finger-Zn knuckle-PHD finger) domain of AF10 senses unmodified H3K27 to regulate DOT1L-mediated methylation of H3K79 (Chen et al. *Mol. Cell* 2015; with the Or Gozani lab). Binding of histone H3 triggers a rearrangement of the PZP module to form an H3(22-27)-accommodating channel with the unmodified H3K27 side chain encased in a compact hydrogen bond acceptor-lined cage. In cells, PZP recognition of H3 is required for H3K79 dimethylation, expression of DOT1L-target genes, and proliferation of DOT1L-addicted leukemic cells.

Research championed in the Davis Allis (Rockefeller University) and Haitao Li (Tsinghua University) labs have shown that engineering of the histone-recognition ADD domain of DNMT3a alters the epigenetic landscape and phenotypic features of mouse embryonic stem cells (Noh et al. *Mol. Cell* 2015).

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.

Structural and functional studies on mouse MORC3 establish it to be a GHKL ATPase that localizes to H3K4me3 marked chromatin (Li et al. *Proc. Natl. Acad. Scis.* 2016). The CW domain uses as aromatic cage to bind H3K4me3 and forms extensive hydrogen bonds with the H3 tails.

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.

Chen, S., Ze, Y., Wilkinson, A., Deshpande, A. J., Sidoli, S., Krajewski, K., Strahl, B. D., Garcia, B. A., Armstrong, S. A., Patel, D. J. and Gozani, O. (2015). The PZP domain of AF10 senses unmodified H3K27 to regulate DOT1L-methylation at H3K79. ***Mol. Cell*** 60, 319-327.

Li, S., Yen, L., Pastor, W.A., Johnston, J. B., Shew, C., Du, J., Wu, H., Liu, W., Ho, J., Stender, B., Clark, A. T., Burlingame, A., Whitelaw, E., Daxinger, L., Patel, D. J. and Jacobsen, S. E. (2016). Mouse MORC3 is an ATP dependent dimer that is involved in gene silencing and localizes to H3K4me3 marked chromatin. ***Proc. Natl. Acad. Scis. USA*** 113: E5108-E5116.

Noh, K. M., Wang, H., Kim, H. R., Wenderski, W., Fang, F., Li, C., Dewell, S., Ferris, A., Hughes, S. H., Zheng, D., Melnick, A, Patel, D. J., Li, H. and Allis, C. D. (2015). Engineering of a histone recognition domain in Dnmt3a alters the epigenetic landscape and phenotypic features of mouse ESCs. ***Mol. Cell*** 59, 89-103.

Oh, Y-S., Gao, P., Lee, K., Ceglia, I., Zhang, X., Ahn, J-H., Chait, B. T., Patel, D. J., Kim, Y. & Greengard, P. (2013). SMARCA3, a chromatin remodeling factor, is required for p11-dependent anti-depressive action. ***Cell*** 152, 831-843.

Xiao, A., Li, H., Shechter, D., Ahn, S. H., Fabrizio, L., Erajument-Bromage, H., Murakami-Ishibe, S., Wang, B., Tempst, P., Hofmann, K., Patel, D. J., Elledge, S. J. & Allis, C. D. (2009). WSTF regulates the DNA damage response of H2A.X via a novel tyrosine kinase activity. ***Nature*** 457, 57-62.

5. Histone Chaperones

The biology of histone proteins encompasses their synthesis in the cytosol, nuclear import and incorporation into nucleosomes, as well as subsequent eviction, redeposition, storage or degradation. Histone chaperones represent a structurally and functionally diverse family of histone-binding proteins that prevent promiscuous interactions of histones before their assembly into chromatin. Current efforts are directed towards an understanding of the mechanism of chaperone-mediated histone shuttling, handover between different chaperone systems, and histone transfer onto and off DNA.

Our studies on histone chaperones have initially focused on the histone-binding domain (HBD) 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. Our studies identify the contributions of key residues on H3.3 and DAXX HBD to chaperone-mediated H3.3 variant recognition specificity (Elsasser et al. *Nature* 2012; with David Allis lab).

The EBV encoded tegument protein BNRF1 is a DAXX-interacting protein required for the establishment of selective viral gene expression during latency. We have solved the structure of the DAXX-interacting domain (DID) of BNRF1 bound to the histone-binding domain (HBD) of DAXX and histone H3.3-H4 (Huang et al. *Nat. Commun*. 2016; with Paul Lieberman, Wistar Institute). The BNRF1 DID contacts DAXX HBD, as well as H3.3 and H4 through four extended loop domains with the BNRF1-DAXX interface responsible for BNRF1 localization to PML-nuclear bodies typically associated with host anti-viral resistance and transcriptional repression. Our findings reveal molecular details of a virus reprogramming of anti-viral histone chaperone to promote viral latency and cellular immortalization.

We have reported on structure-function studies that provide insights into a unique binding mode that enables MCM2, as part of the replicative helicase, to chaperone histone H3-H4 at replication forks. Our structural studies established that two MCM2 HBDs are bound to an H3-H4 tetramer, thereby hijacking interaction sites used by nucleosomal DNA. In addition, our structural studies establish that one each of MCM2 HBD and ASF1 co-chaperone an H3-H4 dimer (Huang et al. *Nat. Struct. Mol. Biol.* 2015; with Anja Groth, University of Copenhagen, Denmark). Functional studies demonstrate that MCM2 HBD is required for MCM2-7 histone chaperone function and normal cell proliferation. In addition, MCM2 can chaperone both new and old canonical histone H3-H4, as well as H3.3 and CENP-A variants.

We have extended our structural studies on the MCM2 HBD chaperone bound to histones H3-H4 to its complex with added Ankyrin repeat domain (ARD) of TONSL, a component of the TONSL–MMS22L homologous recombination (HR) complex (Saredi et al. *Nature* 2016; with Anja Groth lab, Denmark). Our studies identified the TONSL ARD as a reader of histone H4 tails unmethylated at K20 (H4K20me0), which are specific to new histones incorporated during DNA replication and mark post-replicative chromatin until G2/M. Functional studies establish that H4K20me0 recognition is required for TONSL-MMS22L binding to chromatin and accumulation at challenged replication forks and DNA lesions. Together, this reveals a histone reader based mechanism to recognize the post-replicative state, offering a new approach and opportunity to understand DNA repair.

We have also undertaken structure-function studies of histone H3-H4 tetramer maintenance during transcription by chaperone Spt2. Our structural studies demonstrate that hSpt2 HBD is bound to the periphery of the H3-H4 tetramer, mimicking the trajectory of nucleosomal-bound DNA. Functional studies show that contacts between both human and yeast Spt2 HBD with the H3-H4 tetramer are required for the suppression of H3-H4 exchange as measured by H3K56ac and new H3 deposition (Chen et al. *Genes Dev.* 2015; with Amine Nourani lab, University Laval, Quebec).

The CENP-T/-W histone fold complex, as an integral part of the inner kinetochore, is essential for building a proper kinetochore at the centromere in order to direct chromosome segregation during mitosis. We identified SPT16 and SSRP1, subunits of the H2A-H2B histone chaperone FACT, as CENP-W binding partners through a proteomic screen and further showed that the C-terminal region of Spt16 binds to the histone fold region of CENP-T/-W and is competed off by H2A-H2B (Prendergast et al. *Genes Dev*. 2016; with Genevieve Almouzni lab, Institute Curie, Paris). Based on these findings we

propose that FACT acts as a switching chaperone binding H2A-H2B or CENP-T/-W, thereby providing a molecular mechanism underlying constitutive centromere-associated network establishment.

Chen, S., Ruflange, A., Huang, H., Nourani, A. and Patel, D. J. (2015). Structure-function studies of histone H3/H4 tetramer maintenance during transcription by chaperone Spt2. ***Genes Dev*.** 29, 1326-1340.

Elsasser, S. J., Huang, H., Lewis, P. W., Allis, C. D. & Patel, D. J. (2012). DAXX histone chaperone envelops an H3.3/H4 dimer for H3.3-specific recognition. ***Nature*** 491, 560-565.

Huang, H., Deng, Z., Vladimirova, O., Wiedmer, A., Lu, F., Lieberman, P. M. & Patel, D. J. (2016). Structural basis underlying viral hijacking of a histone chaperone complex. ***Nat. Commun.*** 7: 12707.

Huang, H., Stromme, C. B., Saredi, G., Hodl, M., Strandsby, A., Strandsby, A., Gonzalez-Aguilera, C., Chen, S., Groth, A. and Patel, D. J. (2015). A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks. ***Nat. Struct. Mol. Biol.*** 22, 618-626.

Prendergast, L., Muller, S., Liu, Y., Huang, H., Dingli, F., Lowe, D., Vassias, I., Patel, D. J., Sullivan, K. F. and Almouzni, G. (2016). The CENP-T-CENP-W complex is a binding partner of the histone chaperone FACT. ***Genes Dev.*** 30, 1313-1326.

Saredi, G., Huang, H., Hammond, C., Bekker-Jensen, S., Forne, I., Reveron-Gomez, N., Foster, B. M., Mlejnkova, L., Bartke, T., Cejka, P., Mailand, N., Imhof, A., Patel, D. J. and Groth, A. (2016). H4K20me0 marks post-replicative chromatin and recruits the TONSL-MMS22L DNA repair complex. ***Nature*** 534, 714-718.

6. DNA Methylation Mark-mediated Epigenetic Regulation

Methylation of cytosine in the CpG context has pronounced effects on gene expression with DNA methylation patterns established during embryonic development and then faithfully maintained during subsequent somatic cell division. The basic principles underlying the setting up and maintenance of DNA methylation patterns remains an area of intense research, given that perturbation of DNA methylation patterns impacts on a range of human diseases. Current efforts are focused on structure-function studies of writers, readers and erasers of DNA methylation marks and their complexes with unmodified and hemimethylated DNA.

Maintenance DNA Methylation

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. *Scienc*e 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; 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 (UCLA) 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.

Song, J., Rechkoblit, O., Bestor, T. H. & Patel, D. J. (2011). Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. ***Science*** 331,1036-1040.

Song, J., Teplova, M., Ishibe-Murakami, S. & Patel, D. J. (2012). Structural principles underlying DNMT1-mediated DNA methylation. ***Science*** 335, 709-712.

Stroud, H., Do, T., Zhong, X., Feng, S., Johnson, L., Du, J., Patel, D. J. and Jacobsen, S. E. (2014). The roles of non-CG methylation in Arabidopsis. ***Nat. Struct. Mol. Biol.*** 21, 64-72.

Zhong, X., Du, J., Hale, C. J., Feng, S., Gallego-Bartolome, J., Vashisht, A. A., Wohlschlegel, J. A., Patel, D. J. and Jacobsen, S. E. (2014). Molecular mechanism of action of plant DRM *de novo* DNA methyltransferases. ***Cell*** 157, 1050-1060.

Cross-talk Between Histone and DNA Methylation

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, (see Rajakumara et al. *Genes Dev.* 2011) 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.

In earlier studies, we had shown that PRMT5-medited methylation of histone H4R3 recruits DNMT3A, thereby coupling histone and DNA methylation in gene silencing (Zhao et al. *Nat. Struct. Mol. Biol.* 2009; with Stephen Jane, Melbourne)

We have written a structure-function review outlining DNA methylation pathways and their crosstalk with histone methylation (Du et al. *Nat. Rev. Mol. Cell Biol*. 2015).

Du, J., Johnson, L. M., Groth, M., Feng, S., Hale, C. J., Li, S., Vashisht, A. A., Wohlschlegel, J. A., Patel, D. J. and Jacobsen, S. E. (2014). Mechanism of DNA methylation-directed histone methylation by KRYPTONITE. ***Mol. Cell*** 55, 495-504.

Du, J., Johnson, L. M., Jacobsen, S. E. and Patel, D. J. (2015). DNA methylation pathways and their crosstalk with histone methylation. ***Nat. Rev. Mol. Cell Biol*.** 16, 519-532.

Du, J., Zhong, X., Barnatavichute, Y. V., Stroud, H., Feng, S., Caro, E., Vashisht, A. A., Terragni, J., Chin, H. G., Tu, J., Hetzel, J., Wohlschlegel, J. A., Pradhan, S., Patel, D. J. & Jacobsen, S. E. (2012). Dual binding of chromomethylase BAH and chromo domains to H3K9me2-containing nucleosomes in the targeting of DNA methylation. ***Cell*** 151,167-180.

Rajakumara, E., Law, J. A., Shimanshu, D. K., Voigt, P., Johnson, L. M., Reinberg, D., Patel, D. J. & Jacobsen, S. E. (2011). A dual flip out mechanism for 5mC recognition by the *Arabidopsis* SUVH5 SRA domain and its impact on DNA methylation and H3K9 dimethylation *in vivo.* ***Genes Dev*.** 25, 137-152.

Zhao, Q., Rank, G., Tan, Y. T., Li, H., Moritz, R. L., Simpson, R. J., Cerruti, L., Curtis, D. J., Patel, D. J., Allis, C. D., Cunningham, J. M. & Jane, S. M. (2009). PRMT5-mediated methylation of histone H4R3 recruits DNMT3A coupling histone and DNA methylation in gene silencing. ***Nat. Struct. Mol. Biol*.** 16, 304-311.

RNA-directed DNA 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.

Johnson, L. M., Du, J., Chodavarapu, R. K., Feng, S., Zhong, X., Hale, C. J., Marson, G., Segal, D. J., Pellergrini, M., Patel, D. J. and Jacobsen, S. E. (2014). SRA/SET domain proteins link RNA polymerase V binding to DNA methylation. ***Nature*** 507, 124-128.

Law, J. A., Du, J., Hale, C. J., Feng, S., Krajewski, K., Strahl, B. D., Patel, D. J. & Jacobsen, S. E. (2013). SHH1 recruits RNA polymerase-IV to RNA-directed DNA methylation targets. ***Nature*** 498, 385-389.

7. cGAS-cGAMP-STING Innate Immune Response Pathway

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).

Gao, P., Ascano, M., Wu, Y., Barchet, W., Gaffney, B. L., Zillinger, T., Serganov, A., Jones, R. A., Hartmann, G., Tuschl, T. and Patel, D. J. (2013a). Cyclic [G(2',5')pA(3',5')p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. ***Cell*** 153, 1094-1107.

Gao, P., Ascano, M., Zillinger, T., Wang, Y., Dai, P., Serganov, A. A., Gaffney, B. L., Shuman, S., Jones, R., Deng, L., Hartmann, G., Barchet, W., Tuschl, T. and Patel, D.J. (2013b). Structure-function studies of STING activation by c[G(2',5')pA(3',5')p], its linkage isomers and DMXAA. ***Cell*** 154, 748-762.

Gao, P., Zillinger, T., Wang, W., Ascano, M., Dai, P., Hartmann, G., Tuschl, T., Deng, L., Barchet, W. and Patel, D. J. (2014). Binding pocket and lid region substituents render human STING sensitive to mouse-selective drug DMXAA. ***Cell Reports*** 8, 1668-1676.

Wang, Y., Ludwig, J., Schuberth, C., Goldeck, M., Schlee, M., Li, H., Juranek, S., Sheng, G., Micura, R., Tuschl, T., Hartmann, G. & Patel, D. J. (2010). Structural and functional insights into 5'-ppp-RNA pattern recognition by the innate immune receptor RIG-I. ***Nat. Struct. Mol. Biol.*** 17, 781-787.

8. 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* Accelarated 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).

We have written a review outlining details of systems involving sphingolipid transfer proteins that are defined by the GLTP fold (Malinina et al. *Quart. Rev. Biophys*. 2015).

Malinina, L., Malakhova, M. L., Kanack, A. T., Brown, R. E. & Patel, D. J. (2006). The liganding mode of glycolipid transfer protein is controlled by glycolipid acyl structure. ***PLoS Biol*** 4, 1996-2011.

Malinina, L., Malakhova, M. L., Teplov, A., Brown, R. E. & Patel, D. J. (2004). Structural basis for glycosphingolipid transfer specificity. ***Nature*** 430, 1048-1053.

Malinina, L., Simanshu, D. K., Zhai, X., Samygina, V. R., Kamlekar, R., Kenoth, R., Ochoa-Lizarralde, B., Molotkovsky, J. G., Patel, D. J. and Brown, R. E. (2015). Sphingolipid transfer proteins defined by the GLTP-fold. ***Quart. Rev. Biophys*.** 48, 281-322.

Samygina, V., Popov. A. N., Cabo-Bilbao, A., Ochoa-Lizarralde, B., Goni-de-Cerio, F., Zhai, X., Molotkovsky, J. G., Patel, D. J. Brown, R. E., & Malinina, L. (2010). A designer human glycolipid transfer protein with enhanced transfer selectivity for sulfatide. ***Structure*** 19, 1625-1634.

Simanshu, D. K., Kamlekar, R. K., Wijesinghe, D. S., Zou, X., Zhai, X., Mishra, S. K., Molotkovsky, J. G., Malinina, L., Hincliffe, E. H., Chalfant, C. E., Brown, R. E. & Patel, D. J. (2013). Nonvesicular trafficking by ceramide-1-phosphate transfer protein regulates eicosanoid production. ***Nature*** 500, 463-467.

Simanshu, D., Zhai, X., Munch, D., Hofius, D., Markham, J. E., Bielawski, J., Bielawski, A., Slotte, J. P., Malinina, L., Molotkovsky, J. G., Mundy, J. E., Patel, D. J. and Brown, R. E. (2014). Arabidopsis accelerated-cell-death 11 (ACD11) is a ceramide 1-phosphate transfer protein and intermediary regulator of phytoceramide levels. ***Cell Reports*** 6, 388-399.

9. Protein-RNA Complexes on Disease-related Systems

Our group has undertaken x-ray and NMR structural studies on complexes of peptides and proteins bound to their RNA and DNA targets to decipher principles associated with complex formation.

Protein-RNA Complexes

Our group has undertaken x-ray and NMR structural studies on complexes of peptides and proteins bound to their RNA targets to decipher principles associated with complex formation, as part of an effort towards the eventual formulation of a recognition code mediating protein-RNA complex formation. Our studies provide insights into the requirement for the structural integrity of both RNA-binding and dimerization domains of disease-related proteins, as well as their relative orientations, for their post-transcriptional regulatory functions.

These protein-RNA complexes studied by our group include structure-function studies on La autoantigen complexes implicated in autoimmune diseases (Teplova et al. *Mol. Cell* 2006), alternate splicing regulation by MBNL1 (Teplova et al. *Nat. Struct. Mol. Biol*. 2008) and CUGBP1 (Teplova et al. *Structure* 2010) complexes on myotonic dystrophy. We have also studied nucleocytoplasmic export of retroviral genomic RNA (Teplova et al. *Nat. Struct. Mol. Biol.* 2011; with Elisa Izaurralde lab, Max-Planck Institute, Tubingen), and KH domain containing onconeural Nova complexes implicated in POMA syndrome (Teplova et al. *Structure* 2011; with Robert and Jennifer Darnell lab, Rockefeller). In addition, these structure-function studies have been extended to FMRP RGG peptide bound to quadruplex-duplex junctional RNA (Phan et al. *Nat. Struct. Mol. Biol*. 2011; with Robert and Jennifer Darnell; Vasiliyev et al. *Proc. Natl. Acad. Scis. USA* 2015; Alexander Serganov, NYU Medical School), and STAR quaking protein complexes involved in myelination (Teplova et al. *Genes Dev*. 2013; with Thomas Tuschl lab, Rockefeller). More recently we have reported on MazF toxin-RNA and MazF-MazE toxin-antitoxin complexes (Simanshu et al. *Mol. Cell* 2013: with Masayori Inouye lab, Robert Woods Johnson Medical School, NJ) and recognition of distinct RNA motifs by the clustered CCCH zinc fingers of neuronal protein Unkempt involved in the control of a neuronal morphology program (Murn et al *Genes Dev.* 2014; Murn et al. *Nat. Struct. Mol. Biol.* 2016; with Yang Shi lab, Harvard Medical School). More recently, we have investigated the structure and function of the bacterial decapping enzyme NudC (Hofer et al. *Nat. Chem. Biol.* 2016).

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.

Hofer, K., Li, S., Du, J., Abele, F., Frindert, J., Schlothauer, J., Grawenhoff, J., Patel, D. J. & Jaschke, A. (2016). Structure and function of the bacterial decapping enzyme NudC. ***Nat. Chem. Biol.*** 12, 730-734.

Murn, J., Teplova, M., Zarnack, K., Shi, Y. and Patel, D. J. (2016). Recognition of distinct RNA motifs by the clustered CCCH zinc fingers of neuronal protein Unkempt. ***Nat. Struct. Mol. Biol.*** 23, 16-23.

Murn, J., Zarnack, K., Yang, Y. J., Durak, O., Murphy, E. A., Cheloufi, S., Gonzalez, D. M., Teplova, M., Curk, T., Zuber, J., Patel D. J., Ule, J., Luscombe, N. M., Tsai, L. H., Walsh, C. A. and Shi, Y. (2015). Control of a neural morphology program by an RNA-binding zinc finger protein, Unkempt. ***Genes Dev.*** 29, 501-512.

Phan, A. T., Kuryavyi, V., Darnell, J. C., Serganov, A., Majumdar, A., Raslin, T., Polonskaia, A., Chen, C., Clain, D., Darnell, R. B. & Patel, D. J. (2011). Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. ***Nat. Struct. Mol. Biol*.** 18, 796-804.

Serganov, A. & Patel, D. J. (2008). Towards deciphering the principles underlying an mRNA recognition code. ***Curr. Opin. Struct. Biol*.** 18, 120-129.

Simanshu, D. K., Yamaguchi, Y., Park, J-H., Inouye, M. and Patel, D. J. (2013). Structural insights into mRNA recognition by toxin MazF and its regulation by antitoxin MazE in B. subtilis. ***Mol. Cell*** 52, 447-458.

Teplova, M. & Patel, D. J. (2008). Structural insights into RNA recognition by the alternate splicing regulator muscleblind-like MBNL1. ***Nat. Struct. Mol. Biol*.** 15, 1343-1351.

Teplova, M., Hafner, M., Teplov, D., Essig, K., Tuschl, T. and Patel, D. J. (2013). Structure-function studies of STAR family Quaking proteins bound to their *in vivo* RNA target sites. ***Genes Dev.*** 27, 928-940.

Teplova, M., Malinina, L., Darnell, J. C., Song, J., Lu, M., Abagyan, R., Musunuru, K., Teplov, A., Burley, S. K., Darnell, R. B. & Patel, D. J. (2011). Protein-RNA and protein-protein recognition by dual KH1/2 domains of the neuronal splicing factor Nova-1. ***Structure*** 19, 930-944.

Teplova, M., Song, J., Gaw, H. Y., Teplov, V. & Patel, D. J. (2010). Structural insights into RNA recognition by the CUG binding protein 1. ***Structure*** 18, 1364-1367.

Teplova, M., Wohlbold, L., Kim, N. Y., Izaurralde, E. & Patel, D. J. (2011). Structure-function studies of nucleocytoplasmic transport of retroviral genomic RNA by mRNA export factor TAP. ***Nat. Struct. Mol. Biol.*** 18, 990-998.

Teplova, M., Yuan, Y. R., Phan, A. T., Malinina, L., Ilin, S., Teplov, A. & Patel, D. J. (2006). Structural basis for recognition and sequestration of UUU_{OH} 3'-terminii of nascent mRNA polymerase III transcripts by La autoantigen. ***Mol. Cell*** 21, 75-85.

Tian, Y., Simanshu, D. K., Ascano, M., Daiz-Avalos, R., Park, A. Y., Juranek, S. A., Rice, W. J., Yin, Q., Robinson, C. C., Tuschl, T. & Patel, D. J. (2011). Multimeric assembly and biochemical characterization of the Trax-translin endonuclease complex. ***Nat. Struct. Mol. Biol*.** 18, 658-664.

Vasilyev, N., Polonskaia, A., Darnell, J. C., Darnell, R. B., Patel, D. J. and Serganov, A. (2015). Crystal structure reveals specific recognition of a G-quadruplex RNA by a β-turn in the RGG motif of FMRP. ***Proc. Natl. Acad. Scis. USA*.** 112, E5391-E5400.

Protein-DNA Complexes

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*., 2015; championed by the Eric Lai lab).

Dai, Q., Ren, A., Westholm J. O., Patel, D. J. and Lai, E. (2015). Common and distinct DNA-binding and regulatory activities of the BEN-solo transcription factor family. ***Genes Dev*.** 29, 48-62.

Dai, Q., Ren, A., Westholm, J. O., Serganov, A., Patel, D. J. & Lai, E. C. (2013). The BEN domain is a novel sequence-specific DNA binding domain conserved in neural transcriptional repressors. ***Genes Dev.* 27**, 602-614.

10. DNA Damage and Processing by Bypass Polymerases

Our group has been interested in how DNA lesions are accommodated along the DNA duplex as a function of lesion type, chirality and base opposite the lesion site. In addition, we have attempted to understand the principles underlying processing of DNA lesions by bypass polymerases.

Lesion Alignment on DNA

NMR-based structural studies by our group on chiral polycyclic aromatic hydrocarbon (PAH) 'bay' region benzo[a]pyrene-N^2-guanosine (BP-N^2G) and 'fjord' region benzo[c]phenanthrene-N^6-adenosine (BPh-N^6A) 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 Broyde labs, New York University). These are minor groove alignment of BP-N^2G lesion without disruption of the modified base pair (Cosman et al. *Proc. Natl. Acad. Scis. USA* 1992), intercalation of BP-N^2G 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-N^6A 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.

Cosman, M., de los Santos, C., Fiala, R., Hingerty, B. E., Ibanez, V., Luna, E., Harvey, R., Geacintov, N. E., Broyde, S. and Patel, D. J. (1993a). Solution conformation of the (+)-*cis*-anti-[BP]dG adduct in a DNA duplex: Intercalation of the covalently attached benzo[a]pyrenyl ring into the helix and displacement of the modified deoxyguanosine. ***Biochemistry*** 32, 4145-4155.

Cosman, M., de los Santos, C., Fiala, R., Singh, S. B., Hingerty, B. E., Ibanez, V., Margulis, L. A., Live, D., Geacintov, N. E., Broyde, S. and Patel, D. J. (1992). Solution conformation of the major adduct between the carcinogen (+)-*anti*-benzo[a]pyrene diol epoxide and DNA. ***Proc. Natl. Acad. Scis., USA.*** 89, 1914-1918.

Cosman, M., Fiala, R., Hingerty, B. E., Laryea, A., Lee, H., Harvey, R. G., Amin, S., Geacintov, N. E., Broyde, S. and Patel, D. J. (1993b). Solution conformation of the (+)-*trans-anti*-[BPh]dA adduct opposite dT in a DNA duplex: Intercalation of the covalently attached benzo[c]phenanthrene to the 5'-side of the adduct site without disruption of the modified base pair. ***Biochemistry*** 32, 12488-12497.

Geacintov, N. E., Cosman, M., Hingerty, B. E., Amin, S., Broyde, S. & Patel, D. J. (1997). NMR solution structures of stereoisomeric polycyclic aromatic carcinogen-DNA adducts: Principles, patterns and diversity. ***Chem. Res. Toxicol.*** 10, 111-146.

Damage Site Processing by Bypass Polymerases

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 aminofluroene-C8-guanosine (AF-C^8G) lesion (Rechkoblit et al. ***Nat. Struct. Mol. Biol*.** 2010).

Broyde, S., Wang, L., Rechkoblit, O., Geacintov, N. E. & Patel, D. J. (2008). Lesion processing by replicative versus bypass polymerases. ***Trends Biochem. Scis.*** 33, 209-219.

Broyde, S., Wang, L., Zhang, L., Rechkoblit, O., Geacintov, N. E. & Patel, D. J. (2008). DNA adduct structure-function relationships: comparing solution with polymerase structure. ***Chem. Res. Toxicol.*** 21, 45-52.

Rechkoblit, O., Kolbanovskiy, A., Malinina, L., Geacintov, N. E., Broyde, S. & Patel, D. J. (2010). Y-family polymerase-facilitated mechanism of error-free bypass and semi-targeted mutagenic processing of an aromatic amine lesion. ***Nat. Struct. Mol. Biol*.** 17, 379-388.

Rechkoblit, O., Malinina, L., Cheng, Y., Kuryavyi, V., Broyde, S., Geacintov, N. & Patel, D. J. (2006). Stepwise translocation of Dpo4 polymerase during error-free bypass of oxoG lesion. ***PLoS Biology*** 4, 25-42.