

UNIVERSITI PUTRA MALAYSIA

STRUCTURAL AND BIOPHYSICAL CHARACTERISATIONS OF MeCP2 MBD MUTANTS THAT CORRELATE WITH RETT SYNDROME

CHIA JYH YEA

FPSK(M) 2016 47



STRUCTURAL AND BIOPHYSICAL CHARACTERISATIONS OF MeCP2 MBD MUTANTS THAT CORRELATE WITH RETT SYNDROME



By

CHIA JYH YEA

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

January 2016

All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the Degree of Master of Science

STRUCTURAL AND BIOPHYSICAL CHARACTERISATIONS OF MeCP2 MBD MUTANTS THAT CORRELATE WITH RETT SYNDROME

By

CHIA JYH YEA

January 2016

Chairman : Ho Kok Lian, PhD Faculty : Medicine and Health Sciences

Methyl-CpG binding protein (MBD) family consists of Methyl-CpG Binding Protein 2 (MeCP2), Methyl-CpG Binding Domain Protein 1 (MBD1), MBD2, MBD3 and MBD4 where MeCP2 is the prototype of the family. MeCP2 contains several domains: (a) a methyl-CpG binding domain (MBD), (b) a transcriptional repression domain (TRD), (c) two AT hooks and (d) a nuclear localisation signal (NLS). MeCP2 binds to methylated DNA and represses the transcription of the associated genes. Mutations in MECP2 lead to Rett syndrome (RTT), which is characterised by progressive neuro-developmental disorder in early childhood of females. Previous studies revealed that most RTT missense mutations alter the protein conformation and subsequently interferes the methyl-CpG recognition. To understand how the structural changes contribute to RTT, the 3-dimensional structure of these mutants need to be elucidated. Therefore, it is of interest to study the structure of RTT related MBD in complex with methylated DNA using X-ray crystallography and characterised the DNA-protein binding with some biophysical assays. Since more than 50% of the missense mutations occur within the MBD domain. Out of the 8 hot RTT spots within this domain, RTT mutants D97E, A140V, Y141C, P152R and D156E were included in this study. In band shift assay, wild-type MBD complexed with DNA was significantly shifted compared to A140V and D97E while other mutants were not significantly shifted. In SPR, wild-type MBD showed the highest affinity towards the DNA followed by A140V (K_D : 0.28 μ M). Circular dichrosim (CD) analysis revealed that the secondary structures of A140V, Y141C, P152R and D156E are highly similar to wild-type MBD (14.7 % α -helix, 25.2 % β -strand, and 60.1 % turns and unordered) except for D97E which showed 31 % α helix, 6.5 % β -strand, 62.5 % turns and unordered. The complex of MBD^{140V} with methylated DNA was crystallised and diffracted X-ray to 2.2 Å resolution. The cocrystal belongs to monoclinic space group C2, with unit cell parameters of a=78.66 Å, b=53.49 Å, c=62.78 Å, $\alpha = \gamma = 90^{\circ}$ and $\beta = 132.47^{\circ}$. X-ray analysis revealed that the MBD domain was not altered by mutation of Ala-140 to Val (A140V). However, additional water molecules were identified at the DNA-protein contact interface and around the DNA molecule. A narrow minor groove of A/T run was observed as a result of additional bifurcated hydrogen bonds and vertical stacking of bases results from high degree of propeller twist and heavy purine-purine stacking. Two hydration spines were observed running down the wall of the minor groove. Each hydration spine is well

arranged into two shells adopting a zig-zag arrangement. Hence, this finding provides insights for the DNA geometry where the A/T run is geometrically stabilized by extensive water network and is independent of the flanking nucleotide sequence, DNA methylation and the bound MBD domain. The finding explores characteristics of the methylated DNA containing A/T run, which provide the nucleotide sequence preferences to MeCP2. In general, these additional molecular details could provide fundamental knowledge in RTT therapeutic approaches.



Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk Ijazah Master Sains

KAJIAN STRUKTUR DAN BIOFIZIKAL MUTAN MeCP2 MBD YANG BERKAITAN DENGAN SINDROM RETT

Oleh

CHIA JYH YEA

Januari 2016

Pengerusi : Ho Kok Lian, PhD Fakulti : Perubatan dan Sains Kesihatan

Keluarga protein pengikat metil-CpG (MBD) terdiri daripada protein Pengikat Metil-CpG 2 (MeCP2), pengikat metil-CpG 1 (MBD1), MBD2, MBD3 dan MBD4 di mana MeCP2 adalah prototaip keluarga. MeCP2 mengandungi beberapa domain: (a) domain pengikat metil-CpG (MBD), (b) domain penindasan transkripsi (TRD), (c) dua pencangkuk AT dan (d) isyarat penempatan nuklear (NLS). Protein Pengikat Metil-CpG 2 (MeCP2) mengikat DNA bermetil dan menindas transkripsi gen yang berkaitan. Mutasi dalam MECP2 membawa kepada sindrom Rett (RTT), yang mempunyai ciriciri progresif gangguan neuron dalam perkembangan awal kanak-kanak perempuan. Kajian sebelum ini menunjukkan bahawa kebanyakan mutasi missense RTT mengubah komformasi protein dan seterusnya mengganggu pengenalan metil-CpG. Untuk memahami bagaimana perubahan struktur membawa kepada RTT, struktur 3-dimensi mutan perlu dijelaskan. Oleh itu, adalah penting untuk mengkaji struktur RTT berkaitan dengan MBD kompleks dengan DNA bermetil menggunakan sinaran-X penghabluran dan mencirikan dengan beberapa analisis DNA biofizikal metilasi. Lebih daripada 50% daripada mutasi missense berkelompok dalam domain MBD. Daripada 8 lokasi penting RTT dalam domain ini, mutan RTT termasuk D97E, A140V, Y141C, P152R dan D156E telah dimasukkan dalam kajian ini. Dalam essei anjakan jalur, Kompleks DNA-protein MBD jenis liar meranjak dengan ketara berbanding dengan mutan A140V dan D97E manakala anjakan jalur mutan lain tidak ketara. Dalam SPR, MBD jenis liar mempunyai daya afiniti ke arah DNA yang paling tinggi dan diikuti oleh mutan A140V (K_D : 0.28 µM). Analisis dikreisme membulat (CD) menunjukkan bahawa struktur sekundur mutan RTT (A140V, Y141C, P152R dan D156E) adalah menyerupai MBD jenis liar (14.7 % α-heliks, 25.2 % β-helai, dan 60.1 % tidak tersusun) kecuali D97E vang menunjukkan 31 % α -heliks, 6.5 % β -helai, 62.5 % tidak tersusun. Kompleks MBD^{140V} dengan DNA bermetil telah dihablurkan dengan menggunakan 30 % (w/v) PEG 2000 dan hablur telah dibelau oleh sinaran-X kepada resolusi 2.2 Å. Hablur ini tergolong dalam kumpulan ruang monoklinik C2, dengan parameter unit sel = 78.66 Å, b = 53.49 Å, c = 62.78 Å, $\alpha = \gamma = 90^{\circ}$ dan $\beta = 132.47^{\circ}$. Analisis sinaran-X mendedahkan bahawa domain MBD tidak berubah akibat mutasi Ala-140 kepada Val (A140V). Walau bagaimanapun, lebih banyak molekul air telah dikenalpasti di antara muka DNA-protein dan di sekitar molekul DNA. Lurah minor dengan urutan A/T yang berulang telah dikenalpasti dengan tambahan ikatan hidrogen bercabang dan penindihan nukleotida secara menegak yang berpunca daripada pusingan baling-baling

yang berdarjah tinggi dan penindihan purina. Dua spina hidrasi didapati menuruni dinding lurah minor. Setiap spina hidrasi tersusun dalam dua lapisan dengan penyusunan zig-zag. Oleh itu, kajian ini dapat meningkatkan pemahaman dalam geometri DNA di mana urutan A/T yang berulang adalah stabil akibat daripada rangkaian molekul air dan bebas daripada mengapitnya urutan nukleotida, metilasi DNA dan domain MBD yang terikat. Penemuan ini membongkarkan ciri-ciri DNA bermetil yang mengandungi urutan A/T berulang bagi membekalkan kegemaran jujukan nukleotida kepada MeCP2. Secara umum, Informasi molekular yang didapati dapat menawarkan perkembangan dalam permahaman asas dalam terapeutik RTT.



ACKNOWLEDGEMENTS

With God's warmest blessing, there are some people who made my journey easier and I would like to extend my greatest appreciation to those who helped me in completing my journey. First and foremost, I would like to express my heartfelt gratitude to the chairperson of the supervisory committee, Dr. Ho Kok Lian, for his endless support and invaluable guidance throughout this study. He is always in an amiable manner with his friendly gesture. All the gracious personalities of him have motivated me to become a more down to earth and heart-warming person in the near future. In addition, his impressive cognitive ability and intellectual maturity have also thought me the importance of critical and analytical reasoning, not only as a student point of view but also as a researcher.

I would also like to extend my sincere appreciation to the members of the supervisory committee, Prof. Dr. Tan Wen Siang and Assoc. Prof. Dr. Foo Hooi Ling, for their wise opinions and constructive suggestions. Their endurance and boundless enthusiasm in research are some of the characteristics that inspired me in the attempt to become a diligent, persistent and successful researcher. I am also grateful to all my labmates, especially Yoon Kam Yee; for their stimulating discussions, advice, their kind hospitality, friendship and technical assistance. A million thanks goes to all of you who have brightened and amused me in one way or another for making my long exhausting hours in lab a happy and pleasant one. Finally, my special thanks and deepest gratitude are extended to my beloved family members for their understanding, perseverance, support and endless love throughout my study.

Saya mengesahkan bahawa satu Jawatankuasa Peperiksaan Tesis telah berjumpa pada 15 Januari 2016 untuk menjalankan peperiksaan akhir bagi Nurul Shahida binti Hamdan bagi menilai tesis beliau yang bertajuk "Profil, Motif dan Persepsi Pelancong terhadap Langkawi, Malaysia sebagai Destinasi Pelancongan Sukan" mengikut Akta Universiti dan Kolej Universiti 1971 dan Perlembagaan Universiti Putra Malaysia [P.U.(A) 106] 15 Mac 1998. Jawatankuasa tersebut telah memperakukan bahawa calon ini layak dianugerahi ijazah Master Sains

Ahli Jawatankuasa Peperiksaan Tesis adalah seperti berikut:

Kok Lian Yee, PhD Pensyarah Kanan Fakulti Pengajian Pendidikan Universiti Putra Malaysia (Pengerusi)

Soh Kim Geok, PhD Profesor Madya Fakulti Pengajian Pendidikan Universiti Putra Malaysia (Pemeriksa Dalam)

Hamdan Mohd Ali, PhD Pensyarah Kanan Universiti Malaya Malaysia (Pemeriksa Luar)

ZULKARNAIN ZAINAL, PhD Profesor dan Timbalan Dekan Sekolah Pengajian Siswazah Universiti Putra Malaysia

Tarikh: 21 April 2016

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

Ho Kok Lian, PhD

Senior Lecturer Faculty of Medicine and Health Sciences Universiti Putra Malaysia (Chairperson)

Tan Wen Siang, PhD

Professor Faculty of Biotechnology and Biomolecular Sciences Universiti Putra Malaysia (Member)

Foo Hooi Ling, PhD

Associate Professor Faculty of Biotechnology and Biomolecular Sciences Universiti Putra Malaysia (Member)

BUJANG KIM HUAT, PhD Professor and Dean

School of Graduate Studies Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work
- quotations, illustrations and citations have been duly referenced
- the thesis has not been submitted previously or comcurrently for any other degree at any institutions
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be owned from supervisor and deputy vice –chancellor (Research and innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature:	Date:	
e -		

Name and Matric No.: Chia Jyh Yea, GS30076

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: Name of Chairman of Supervisory Committee:	Dr.Ho Kok Lian,
Signature:	
Name of Member	
of Supervisory	Professor
Committee:	Dr.Tan Wen Siang
Signature:	
Name of Member	
of Supervisory	Associate Professor
Committee:	Dr.Foo Hooi Ling

TABLE OF CONTENTS

			Page
ABSTRACK	ζ.		i
ABSTRACT	7		iii
ACKNOWI	EDGEME	INTS	V
APPROVAL			vi
DECLARA	- FION		viii
LIST OF TA	ABLES		xiii
LIST OF FI	GURES		xiv
LIST OF A	BBREVIA	LIONS	xvi
LIST OF AN	MINO ACI	DS ABBREVIATIONS	xix
CHAPTER			
1		DUCTION	1
	1.1	General Introduction	1
	1.2	Objectives	3
2	LITERA	TURE REVIEW	4
	2.1	DNA methylation and epigenetics	4
	2.2	DNA methylation represses gene expression	4
	2.3	Methyl CpG Binding Protein Family	5
		2.3.1 Identification of Methyl-CpG binding protein	5
		2.3.2 MBD1	6
		2.3.3 MBD2 and MBD3	6
		2.3.4 MBD4	7
	2.4	MeCP2	8
		2.4.1 Methyl CpG binding protein 2 (MeCP2)	8
		2.4.2 MeCP2 binding partners	8
		2.4.3 MeCP2 and Rett syndrome	8
	2.5	Structure of methyl-CpG binding domains	9
	2.6	Protein chromatography	13
		2.6.1 Immobilised Metal Affinity Chromatography	
		(IMAC)	13
		2.6.2 Sephacryl S-200 gel filtration	13
	2.7	Biophysical characterisations	14
		2.7.1 Circular dichroism	14
		2.7.2 Electrophoretic mobility shift assay (EMSA)	15
	• •	2.7.3 Surface Plasmon Resonance (SPR)	16
	2.8	Protein Crystallography	16
		2.8.1 Method of Crystallisation	17
		2.8.2 Cryo-crystallography	20
		2.8.5 A-ray dillifaction 2.8.4 Data collection and processing strategy	20
		2.8.4 Data collection and processing strategy 2.8.5 Molecular replacement	22
		2.0.3 MORCULAR replacement 2.8.6 Refinement and final validation	23 23
		2.0.0 Remement and illiar valuation	23
3	MATERI	IALS AND METHODS	25
	3.1	Materials	25

3.2	3.2 Construction of mutants		
3.2.1 Plasmid extraction		26	
3.2.2 Site-directed mutagenesis		26	
	3.2.3 Preparation of Competent cells		
	3.2.4 Transformation	28	
3.3	Protein expression and purification	28	
	3.3.1 SDS-PAGE	29	
3.4	Preparation of DNA double helix	30	
3.5	Preparation of DNA-protein complex	30	
3.6	Biophysical characterisations	30	
	3.6.1 Circular dichroism	30	
	3.6.2 Electrophoretic mobility shift assay (EMS)	A) 30	
	3.6.3 Surface plasmon resonance (SPR)	31	
3.7	DNA-protein co-crystallisation	31	
	3.7.1 Crystallisation method, conditions and cry	0-	
	protectant soaking	31	
	3.7.2 Data collection and processing	32	
	3.7.3 Molecular replacement, model building a	nd	
	refinement	32	
RESULTS	AND DISCUSSION	33	
4.1	Construction of MBD mutants	33	
4.2	Expression of wild-type MBD and MBD mutants	36	
4 <mark>.</mark> 3	Purification of wild-type MBD and MBD mutants	37	
4.4	Characterisation of MBD mutants and in complex w	vith	
	methylated DNA	39	
	4.4.1 Circular dichroism (CD)	39	
	4.4.2 Electrophoretic Mobility Shift assay (EMS	SA) 41	
	4.4.3 Surface Plasmon resonance (SPR)	42	
4.5	Crystallisation	45	
	4.5.1 Crystallisation of MBD mutant D97E in co	omplex	
	with methylated DNA	47	
	4.5.2 Crystallisation of MBD mutant A140V in		
	complex with methylated DNA	48	
4.6	Data collection, processing and integration	49	
	4.6.1 Molecular replacement	52	
	4.6.2 Model building and refinement	52	
4.7	Structural analysis of the MeCP2 MBD ^{140V} –methyl	ated	
	DNA complex	56	
	4.7.1 Structural analysis of MBD bound methyla	ated	
	DNA	58	
	4.7.2 Spine of hydration in the A/T run minor gr	oove 60	
4.8	DNA Geometry	65	
	4.8.1 Geometrical description of dinucleotides s	teps 65	
	4.8.2 Overall DNA geometry	66	
	4.8.3 Bend angle of DNA geometry	69	
	4.8.4 High degree of propeller twist	70	
	Conclusion	15	
5.1	Conclusion Future perspective	13	
5.2	ruture perspective	13	

4

5

 \bigcirc

REFERENCES APPENDICES BIODATA OF STUDENT LIST OF PUBLICATIONS

 (\mathbf{C})



LIST OF TABLES

Table		Page
3.1	Common buffers and chemical compositions .	25
3.2	Primers used to construct MBD mutants.	27
3.3	Thermal profile used in QuikChange site-directed mutagenesis.	27
4.1	Secondary structure analysis for wild-type MBD and MBD mutants.	40
4.2	Data collection strategies.	49
4.3	Crystallographic statistics for MBD in complex with methylated DNA.	50
4.4	Thermal factor of TLS motion groups.	53
4.5	Refinement statistics.	54
4.6	Secondary structure of X-ray MBD ^{140V} –methylated DNA complex structure analysed with PROMOTIF.	57
4.7	MBD β-turns.	58
4.8	Structured water molecules that bridge directly between protein and DNA sugar-phosphate backbone.	60
4.9	Water bridges at the AT run.	64
4.10	Comparison of the minor and major groove width of the DNA fragment of used in this study with at centres of dodecamer from Dickerson-Drew (PDB:1BNA).	of 67
1 11	Roll and Tilt of local base pair step parameter	69
7.11	Non and Theor local base pair step parameter.	09
4. 12	Propeller twist of local base pair parameters.	71

6

LIST OF FIGURES

Figure		Page
2.1	MeCP2 function in gene regulation and chromatin remodeling.	5
2.2	The methyl-CpG binding protein family.	5
2.3	Hydrolytic deamination.	7
2.4	Eight hot mutation spots within the MBD domain.	9
2.5	Overlay of MBD domains among MBD family members.	12
2.6	Immobilised metal affinity chromatography.	13
2.7	Size exclusion chromatography.	14
2.8	Far UV CD spectra associated with various type of secondary structures.	15
2.9	Schematic illustration of protein crystallisation phase diagram.	17
2.10	Hanging drop vapour diffusion.	18
2.11	Sitting drop vapour diffusion.	19
2.12	Bragg's Law.	21
2.13	Sphere of reflection or Ewald sphere construction.	22
4.1	Side-directed mutagenesis of MBD.	35
4.2	SDS-PAGE [15% (w/v)] analysis of expressed wild-type MBD.	36
4.3	SDS-PAGE [15% (w/v)] of expressed wild-type MBD and MBD mutan	ts.36
4.4	SDS-PAGE [15% (w/v)] analysis of protein fractions separated from MI mutant A140V by IMAC affinity column.	3D 37
4. 5	SDS-PAGE [15 % (w/v)] analysis of protein fraction separated from poor IMAC active fraction of MBD mutant A140V.	led 38
4.6	Circular dichroism spectrum of wild-type MBD and MBD mutants.	40
4.7	Band shift assay.	42
4.8	Sensorgram of Surface plasmon resonance.	43
4.9	Sensorgram of the binding of wild-type MBD and MBD mutants to Biotimethylated DNA.	n- 44

 \bigcirc

4. 10	Comparison of binding levels of the wild-type MBD and the MBD mutants Biotin-methylated DNA.	s to 45
4.11	Multiple sequence alignment of MBD domain of MBD proteins.	46
4.12	Identity and similarity of MBD domain of MBD proteins.	47
4. 13	Crystallisation of D97E mutant in complex with methylated DNA.	48
4. 14	Crystallisation of A140V mutant in complex with methylated DNA.	49
4. 15	Data collection of crystal DNA-MBD ^{140V} .	51
4. 16	Model building of DNA-MBD ^{140V} .	52
4. 17	TLS motion groups of DNA-MBD ^{140V} model.	54
4. 18	Analysis of stereochemical properties of DNA-MBD ^{140V} X-ray structure.	55
4. 19	Overall X-ray structure of MBD ^{140V} in complex with methylated DNA at 2.18 Å.	56
4. 20	E137 is stabilised by water molecules and amide groups.	59
4. 21	Hydrogen bonds between O2' of m5C8 with W63 and W64.	59
4. 22	Overall solvent structure of MBD140V in complex with methylated DNA.	62
4. 23	Bridging system in AT run.	63
4.24	Geometrical description of DNA parameters.	65
4. 25	Graphical output of a Curves+ analysis for DNA.	66
4. 26	Overlay of idealised B-DNA (Blue) and X-ray DNA (Cyan) with A/T run highlighted in red.	68
4. 27	Superimposed of A/T run of MBD ^{140V} with HMGA-1.	68
4. 28	Stereo view of the superposition of the A/T run.	70
4. 29	High degree of propeller twist at AT run base pair.	72

LIST OF ABBREVIATIONS

α	Alpha
Å	Ångström
β	Beta
μg	microgram (10 ⁻⁶ g)
μL	microliter (10 ⁻⁶ L)
μΜ	micromolar (10 ⁻⁶ M)
BDNF	Brain-derived neurotrophic factor
bp	Basepair
BSA	bovine serum albumin
ссс	covalently closed circular
CCP4	Collaborative Computational Project Number 4
C-terminal	carboxyl terminal
DNA	Deoxy-ribonucleic acid
dNTP	deoxynucleoside triphosphate
dsDNA	double stranded DNA
DTT	1,4-dithiothreitol
EDTA	ethylene diamine tetraacetic acid
EMSA	Electrophoretic mobility shift assay
FPLC	fast protein liquid chromatography
HDAC	Histone deacetylase
HPLC	high performance liquid chromatography
IMAC	immobilised metal ion affinity chromatography
IPTG	isopropyl-β-d-thiogalactopyranoside

K	Kelvin
Kb	kilobase
K _d	dissociation constant
kDA	kilo Dalton
LB	Luria broth
m5C	5'methyl cytosine
MBD	Methyl-CpG Binding Domain
MBD1	Methyl-CpG Binding Domain Protein 1
MBD2	Methyl-CpG Binding Domain Protein 2
MBD3	Methyl-CpG Binding Domain Protein 3
MBD4	Methyl-CpG Binding Domain Protein 4
MeCP2	Methyl-CpG Binding Protein 2
MBP	Methyl-CpG Binding Protein
Mg	milligram (10 ⁻³ g)
Min	minute
NDB	Nucleic Acid Database
NLS	Nuclear localisation Signal
nM	nanomolar (10 ⁻⁹ M)
NTA	Nitrilotriacetic acid
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PCM1	Protein containing MBD 1
PCR	Polymerase chain reaction
PDB	Protein Database Bank
PEG	Polyethylene glycol

pI	Isoelectric point
RMSD	root mean square deviation
rpm	revolutions per minute
RTT	Rett Syndrome
8	Second
SDS	sodium dodecyl sulphate
SEC	size exclusion chromatography
TBE	Tris-buffered EDTA solution
TEMED	Tetramethyl ethylenediamine
TRD	Transcriptional repression domain
U	Unit
UV	Ultraviolet
v	Volt
V_{m}	Matthew's coefficient
v/v	volume/volume
w/v	weight/volume
x g	centrifugal force

 \bigcirc

LIST OF AMINO ACIDS ABBREVIATIONS

	One letter code	Three letter code
Alanine	А	Ala
Arginine	R	Arg
Asparagine	Ν	Asn
Aspartic acid	D	Asp
Cysteine	С	Cys
Glutamic acid	Е	Glu
Glutamine	Q	Gln
Glycine	G	Gly
Histidine	Н	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	М	Met
Phenylalanine	F	Phe
Proline	Р	Pro
Serine	S	Ser
Threonine	Т	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

 \bigcirc

CHAPTER 1

INTRODUCTION

1.1 General Introduction

DNA methylation is an epigenetic signal that affects gene regulation, genomic stability and chromatin structure in mammalian cells (Bird, 2002; Du *et al.*, 2015). In most cases, this signal can be read by a family of proteins that contains a common methyl-CpG binding domain (MBD) (Hendrich and Bird, 1998). To date, five family members, namely MeCP2, MBD1, MBD2, MBD3 and MBD4 have been identified, in which, MeCP2 is the prototype of this family (Hendrich and Bird, 1998). MeCP2, MBD1 and MBD2 are able to recruit co-repressor complexes that can inhibit transcription in concert with chromatin modifiers (Scarsdale *et al.*, 2011). Mammalian MBD3 does not bind to methylated DNA *in vitro* and *in vivo* due to replacement of amino acid (K43H and Y47F) which is critical for DNA binding (Fraga *et al.*, 2003). MBD4 contains a thymine DNAglycosylase at the C-terminal region that can repair G-T mismatches via hydrolytic deamination [(Refer Section 2.3.4 and Figure 2.3) (Neddermann *et al.*, 1996)].

MeCP2 is a transcriptional repressor that contains several domains: (a) a methyl-CpG binding domain (MBD), (b) a transcriptional repression domain (TRD), (c) AT hooks and (d) a nuclear localisation signal (NLS) (Lewis et al., 1992; Nan and Bird, 2001; Xu and Pozzo-Miller, 2013). MBD domain of MeCP2 is able to recognise methyl-CpG containing DNA. TRD domain involves in recruitment of transcriptional co-repressors such as mSin3A and histone deacetylases (HDACs) (Bienvenu et al., 2000). Two putative NLSs facilitate nuclear localisation, which targets the protein into the cell nucleus (Weaving et al., 2004) and the AT hooks are believed to interact with AT rich region of the DNA (Klose et al., 2005). In addition, the AT hooks of MeCP2 bearing amino acid sequences ¹⁸⁵GRGRGRP¹⁹¹ and ²⁶⁵PKKRGRKP²⁷² (superscript indicates amino acid number) which are highly similar to the high mobility group with the AT hook I chromosomal protein (HMCG-I) that is capable to bind to the minor groove of the AT stretches (A/T run) of DNA and functionally (Aravind and Landsman, 1998; Lewis et al., 1992; Nan et al., 1993; Reeves and Nissen, 1990). Baker et al. (2013) demonstrated that the disruption of second conserved AT hook at the C-terminal region of MeCP2 by truncation at R270X of MeCP2 led to failure in chromatin compaction and localization of pericentric heterochromatin domain of α -thalassemia mental retardation syndrome X-linked (ATRX); a chromatin remodelling protein, with MeCP2, and caused the R270X mice to exhibit Rett syndrome (RTT) phenotypes which is similar to MeCP2 knock-out mice (Baker et al., 2013; Xu and Pozzo-Miller, 2013).

The MBD domain of MeCP2 is able to recognise single methyl-CpG dinucleotide (Lewis *et al.*, 1992). The MBD domain alone is ample for the methylated DNA binding and mutations in the MBD domain intercept its binding to methylated sequence (Baubec *et al.*, 2013; Yusufzai and Wolffe, 2000). MeCP2 mutation causes Rett syndrome (RTT); a progressive neurodevelopmental disorder in early childhood, which leads to mental

retardation in females, with a prevalence of 1 in 10,000-15,000 female births (Hagberg, 1985). RTT is caused by an X-linked mutation dominant inheritance with normally lethality in males due to severe encephalopathy (Bianciardi et al., 2015; Bienvenu et al., 2000; Zhao et al., 2015). Studies revealed that most RTT related missense mutations alter the structure of the MBD domain and subsequently interrupt DNA recognition properties (Kriaucionis and Bird, 2003; Kucukkal et al., 2015). Klose et al. (2005) showed that an A/T run adjacent to the methyl-CpG is required to enhance the MeCP2 binding (Klose et al., 2005). Identified endogenous MeCP2 targeting genes such as brain derived neurotropic factor (BDNF) promoter region contains high occurrences of A/T runs closed to the methyl-CpGs (Chen et al., 2003; Martinowich et al., 2003). The A/T run in the methylated DNA facilitated the co-crystallisation of MeCP2 MBD domain in complex with methylated DNA used in this study. Due to the presence of AT hooks in MeCP2 and the requirement of A/T run for maximal binding, it has been speculated that the A/T run could interact with the AT hooks of the MeCP2. However, the characteristics of A/T run which provide specificity for the MeCP2 to recognise the methyl groups remained unclear. Therefore, it is of interest to elucidate the 3-dimensional structure of RTT mutants in order to understand how the structural changes contribute to RTT and the A/T run characteristics with MBD domain bound to its adjacent methyl-CpG dinucleotide.

X-ray analysis of previous report on a MeCP2 MBD-methylated DNA complex revealed that only a few residues are involved in direct contact with the DNA bases (Ho et al., 2008). The methyl groups are recognised by the Arginine fingers of R111 and R133 while D121 is critical in maintaining the unique hydration pattern at the DNA-protein interface. The unique water molecules distribution pattern is crucial to mediate methyl group recognition (Ho et al., 2008). RTT mutations within the MBD domain of MeCP2, however, are believed to alter the 3-dimensional structure of the protein and subsequently affects DNA binding. Several critical mutations such as R111G, R133C, T158M and D121G, which close to the DNA-protein contact region have been investigated (Free et al., 2001; Meehan et al., 1992; Nan et al., 1993; Yusufzai and Wolffe, 2000). In this study, other RTT mutations which are located distance from the DNA-protein contact region have been studied. According to Wakefield and colleagues (1999), missense mutations found in RTT usually do not specifically interrupt DNA recognition but may result in structural changes in the domain (Bianciardi et al., 2015; Wakefield et al., 1999). In order to further investigate the details on the structural changes and molecular functional role, several mutants (D97E, A140V, Y141C, P152R and D156E) were constructed and the DNA-MBD interactions were characterised with various biophysical assays. A co-crystal structure was also elucidated, in which, more molecular details about the DNA-protein complex have been revealed compared with previous reported structure. In addition, the atomic details of the DNA geometry of the MBD bound A/T run are highlighted in comparison with the A/T run of the free DNA double helices. In general, these additional molecular details could provide fundamental knowledge in RTT therapeutic approaches.

1.2 Objectives

The general objective of this study was to explore and understand the atomic details of Rett mutants in complex with methylated DNA. The specific objectives were:

- 1. To construct MBD mutants
- 2. To characterise the interactions of MBD mutants and methylated DNA
- 3. To crystallise the MBD mutants in complex with methylated DNA
- 4. To solve the X-ray structure of MBD mutants in complex with methylated DNA



REFERENCES

- Abrescia N.G., Thompson A., Huynh-Dinh T., Subirana J.A. (2002). Crystal structure of an antiparallel DNA fragment with Hoogsteen base pairing. *Proceedings of the National Academy of Sciences of the United States of America* 99: 2806-2811
- Amir R., Dahle E.J., Toriolo D., Zoghbi H.Y. (2000). Candidate gene analysis in Rett syndrome and the identification of 21 SNPs in Xq. American Journal of Medical Genetics 90: 69-71
- Amir R.E., Van den Veyver I.B., Wan M., Tran C.Q., Francke U., Zoghbi H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics* 23: 185-188
- Aravind L., Landsman D. (1998). AT-hook motifs identified in a wide variety of DNAbinding proteins. *Nucleic Acids Research* 26: 4413-4421
- Baker S.A., Chen L., Wilkins A.D., Yu P., Lichtarge O., Zoghbi H.Y. (2013). An AThook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. *Cell* 152: 984-996
- Bhattacharya S.K., Ramchandani S., Cervoni N., Szyf M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* 397: 579-583
- Bienvenu T., Carrie A., de Roux N., Vinet M.C., Jonveaux P., Couvert P., Villard L., Arzimanoglou A., Beldjord C., Fontes M., Tardieu M., Chelly J. (2000).
 MECP2 mutations account for most cases of typical forms of Rett syndrome. *Human Molecular Genetics* 9: 1377-1384
- Bird A. (2002). DNA methylation patterns and epigenetic memory. *Genes & Development* 16: 6-21
- Bird A. (2008). The methyl-CpG-binding protein MeCP2 and neurological disease. Biochemical Society Transactions 36: 575-583
- Blanchet C., Pasi M., Zakrzewska K., Lavery R. (2011). CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. *Nucleic Acids Research* 39: W68-73
- Block H., Maertens B., Spriestersbach A., Brinker N., Kubicek J., Fabis R., Labahn J., Schafer F. (2009). Immobilized-metal affinity chromatography (IMAC): a review. *Methods in Enzymology* 463: 439-473
- Brigham-Burke M., Edwards J.R., O'Shannessy D.J. (1992). Detection of receptor-ligand interactions using surface plasmon resonance: model studies employing the HIV-1 gp120/CD4 interaction. *Analytical Biochemistry* 205: 125-131

- Chayen N.E. (1998). Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Crystallographica. Section D, Biological Crystallography* 54: 8-15
- Chayen N.E., Saridakis E. (2002). Protein crystallization for genomics: towards highthroughput optimization techniques. *Acta Crystallographica. Section D, Biological Crystallography* 58: 921-927
- Chen W.G., Chang Q., Lin Y., Meissner A., West A.E., Griffith E.C., Jaenisch R., Greenberg M.E. (2003). Derepression of BDNF transcription involves calciumdependent phosphorylation of MeCP2. *Science (New York, N.Y.)* 302: 885-889
- Choi W.I., Jeon B.N., Yoon J.H., Koh D.I., Kim M.H., Yu M.Y., Lee K.M., Kim Y., Kim K., Hur S.S., Lee C.E., Kim K.S., Hur M.W. (2013). The protooncoprotein FBI-1 interacts with MBD3 to recruit the Mi-2/NuRD-HDAC complex and BCoR and to silence p21WAF/CDKN1A by DNA methylation. *Nucleic Acids Research* 41: 6403-6420
- Cramer J.M., Scarsdale J.N., Walavalkar N.M., Buchwald W.A., Ginder G.D., Williams D.C., Jr. (2014). Probing the dynamic distribution of bound states for methylcytosine-binding domains on DNA. *The Journal of Biological Chemistry* 289: 1294-1302
- Crook J.M., Dunn N.R., Colman A. (2006). Repressed by a NuRD. *Nature Cell Biology* 8: 212-214
- Cross S.H., Meehan R.R., Nan X., Bird A. (1997). A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nature Genetics* 16: 256-259
- Dickerson R.E. (1998). DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Research* 26: 1906-1926
- Du Q., Luu P.L., Stirzaker, C. Clark S.J. (2015). Methyl-CpG-binding domain proteins: Readers of the Epigenome. *Epigenomics* 7: 1051-1073
- El Hassan M.A., Calladine C.R. (1996). Propeller-twisting of base-pairs and the conformational mobility of dinucleotide steps in DNA. *Journal of Molecular Biology* 259: 95-103
- Emsley P., Cowtan K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica. Section D, Biological Crystallography* 60: 2126-2132
- Evans P. (2006). Scaling and assessment of data quality. *Acta Crystallographica. Section* D, Biological Crystallography 62: 72-82
- Fonfria-Subiros E., Acosta-Reyes F., Saperas N., Pous J., Subirana J.A., Campos J.L. (2012). Crystal structure of a complex of DNA with one AT-hook of HMGA1. *PloS One* 7: e37120

- Fraga M.F., Ballestar E., Montoya G., Taysavang P., Wade P.A., Esteller M. (2003). The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties. *Nucleic Acids Research* 31: 1765-1774
- Fratini A.V., Kopka M.L., Drew H.R., Dickerson R.E. (1982). Reversible bending and helix geometry in a B-DNA dodecamer: CGCGAATTBrCGCG. *The Journal* of Biological Chemistry 257: 14686-14707
- Free A., Wakefield R.I., Smith B.O., Dryden D.T., Barlow P.N., Bird A.P. (2001). DNA recognition by the methyl-CpG binding domain of MeCP2. *The Journal of Biological Chemistry* 276: 3353-3360
- Fried M., Crothers D.M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Research* 9: 6505-6525
- Fujita N., Takebayashi S., Okumura K., Kudo S., Chiba T., Saya H., Nakao M. (1999). Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms. *Molecular and Cellular Biology* 19: 6415-6426
- Fujita N., Watanabe S., Ichimura T., Ohkuma Y., Chiba T., Saya H., Nakao M. (2003a). MCAF mediates MBD1-dependent transcriptional repression. *Molecular and Cellular Biology* 23: 2834-2843
- Fujita N., Watanabe S., Ichimura T., Tsuruzoe S., Shinkai Y., Tachibana M., Chiba T., Nakao M. (2003b). Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression. *The Journal of Biological Chemistry* 278: 24132-24138
- Garman E.F., Owen R.L. (2006). Cryocooling and radiation damage in macromolecular crystallography. Acta Crystallographica. Section D, Biological Crystallography 62: 32-47
- Ghosh R.P., Horowitz-Scherer R.A., Nikitina T., Gierasch L.M., Woodcock C.L. (2008). Rett syndrome-causing mutations in human MeCP2 result in diverse structural changes that impact folding and DNA interactions. *The Journal of Biological Chemistry* 283: 20523-20534
- Garcia-Ruiz J.M. (2003). Counterdiffusion methods for macromolecular crystallization. *Methods in Enzymology* 368: 130-154
- Greenfield N.J. (1996). Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Analytical Biochemistry* 235: 1-10
- Haff L.A. (1978). Fractionation of water-insoluble protein using Sephacryl S-200 in formamide. *Preparative Biochemistry* 8: 99-112
- Hagberg B. (1985). Rett's syndrome: prevalence and impact on progressive severe mental retardation in girls. *Acta Paediatrica Scandinavica* 74: 405-408

- Harker D. (1956). X-ray diffraction applied to crystalline proteins. Advances in Biological and Medical Physics 4: 1-22
- Hashimoto H., Zhang X., Cheng X. (2012). Excision of thymine and 5hydroxymethyluracil by the MBD4 DNA glycosylase domain: structural basis and implications for active DNA demethylation. *Nucleic Acids Research* 40: 8276-8284
- Hendrich B., Bird A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Molecular and Cellular Biology* 18: 6538-6547
- Hendrich B., Bird A. (2000). Mammalian methyltransferases and methyl-CpG-binding domains: proteins involved in DNA methylation. *Current Topics in Microbiology and Immunology* 249: 55-74
- Hendrich B., Hardeland U., Ng H.H., Jiricny J., Bird A. (1999). The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* 401: 301-304
- Hermann A., Gowher H., Jeltsch A. (2004). Biochemistry and biology of mammalian DNA methyltransferases. *Cellular and Molecular Life Sciences : CMLS* 61: 2571-2587
- Ho K.L., McNae I.W., Schmiedeberg L., Klose R.J., Bird A.P., Walkinshaw M.D. (2008). MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Molecular Cell* 29: 525-531
- Horowitz S., Trievel R.C. (2012). Carbon-oxygen hydrogen bonding in biological structure and function. *The Journal of Biological Chemistry* 287: 41576-41582
- Hutchinson E.G., Thornton J.M. (1996). PROMOTIF--a program to identify and analyze structural motifs in proteins. *Protein science* : A Publication of the Protein Society 5: 212-220
- Jeffery L., Nakielny S. (2004). Components of the DNA methylation system of chromatin control are RNA-binding proteins. *The Journal of Biological Chemistry* 279: 49479-49487
- Jentarra G.M., Olfers S.L., Rice S.G., Srivastava N., Homanics G.E., Blue M., Naidu S., Narayanan V. (2010). Abnormalities of cell packing density and dendritic complexity in the MeCP2 A140V mouse model of Rett syndrome/X-linked mental retardation. *BMC Neuroscience* 11: 19
- Jing M., Bowser M.T. (2011). Methods for measuring aptamer-protein equilibria: A review. *Analytica Chimica Acta* 686: 9-18
- Jones P.L., Veenstra G.J., Wade P.A., Vermaak D., Kass S.U., Landsberger N., Strouboulis J., Wolffe A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genetics* 19: 187-191

- Jorgensen H.F., Ben-Porath I., Bird A.P. (2004). Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. *Molecular and Cellular Biology* 24: 3387-3395
- Kabsch W. (2010). XDS. Acta Crystallographica. Section D, Biological Crystallography 66: 125-132
- Kelly S.M., Jess T.J., Price N.C. (2005). How to study proteins by circular dichroism. *Biochimica et Biophysica Acta* 1751: 119-139
- Khrustalev V.V., Barkovsky E.V., Khrustaleva T.A. (2014). The influence of flanking secondary structures on amino Acid content and typical lengths of 3/10 helices. *International Journal of Proteomics* 2014: 360230
- Klose R.J., Sarraf S.A., Schmiedeberg L., McDermott S.M., Stancheva I., Bird A.P. (2005). DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Molecular Cell* 19: 667-678
- Kriaucionis S., Bird A. (2003). DNA methylation and Rett syndrome. *Human Molecular Genetics* 12 Spec No 2: R221-227
- Kucukkal T.G., Yang Y., Uvarov O., Cao W., Alexov E. (2015). Impact of Rett syndrome mutations on MeCP2 MBD stability. *Biochemistry* 54: 6357-6368
- Laemmli U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-685
- Laskowski R.A., Rullmannn J.A., MacArthur M.W., Kaptein R., Thornton J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR* 8: 477-486
- Leslie A.G. (2006). The integration of macromolecular diffraction data. Acta Crystallographica. Section D, Biological Crystallography 62: 48-57
- Lewis J.D., Meehan R.R., Henzel W.J., Maurer-Fogy I., Jeppesen P., Klein F., Bird A. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* 69: 905-914
- Li T., Jin Y., Vershon A.K., Wolberger C. (1998). Crystal structure of the MATa1/MATalpha2 homeodomain heterodimer in complex with DNA containing an A-tract. *Nucleic Acids Research* 26: 5707-5718
- Lu X.J., Olson W.K. (2003). 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Research* 31: 5108-5121
- Ma L.Y., Wu C., Jin Y., Gao M., Li G.H., Turner D., Shen J.X., Zhang S.J., Narayanan V., Jentarra G., Wu J. (2014). Electrophysiological phenotypes of MeCP2 A140V mutant mouse model. CNS Neuroscience & Therapeutics 20: 420-428

- Magdinier F., Wolffe A.P. (2001). Selective association of the methyl-CpG binding protein MBD2 with the silent p14/p16 locus in human neoplasia. *Proceedings* of the National Academy of Sciences of the United States of America 98: 4990-4995
- Marston F.A. (1986). The purification of eukaryotic polypeptides synthesized in Escherichia coli. *The Biochemical Journal* 240: 1-12
- Martinowich K., Hattori D., Wu H., Fouse S., He F., Hu Y., Fan G., Sun Y.E. (2003). DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science (New York, N.Y.)* 302: 890-893
- Matthews B.W. (1968). Solvent content of protein crystals. *Journal of Molecular Biology* 33: 491-497
- Mcpherson A. (2004). Introduction to protein crystallization. *Methods (san diego, calif.)* 34: 254-265
- Meehan R., Lewis J., Cross S., Nan X., Jeppesen P., Bird A. (1992). Transcriptional repression by methylation of CpG. *Journal of Cell Science. Supplement* 16: 9-14
- Millar C.B., Guy J., Sansom O.J., Selfridge J., MacDougall E., Hendrich B., Keightley P.D., Bishop S.M., Clarke A.R., Bird A. (2002). Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science (New York, N.Y.)* 297: 403-405
- Minor W., Cymborowski M., Otwinowski Z., Chruszcz M. (2006). HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallographica. Section D, Biological Crystallography* 62: 859-866
- Murshudov G.N., Vagin A.A., Dodson E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica. Section D, Biological Crystallography* 53: 240-255
- Nan X., Bird A. (2001). The biological functions of the methyl-CpG-binding protein MeCP2 and its implication in Rett syndrome. *Brain & Development* 23 Suppl 1: S32-37
- Nan X., Campoy F.J., Bird A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* 88: 471-481
- Nan X., Cross S., Bird A. (1998a). Gene silencing by methyl-CpG-binding proteins. *Novartis Foundation Symposium* 214: 6-16; discussion 16-21, 46-50
- Nan X., Hou J., Maclean A., Nasir J., Lafuente M.J., Shu X., Kriaucionis S., Bird A. (2007). Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proceedings of the National Academy of Sciences of the United States of America* 104: 2709-2714

- Nan X., Meehan R.R., Bird A. (1993). Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Research* 21: 4886-4892
- Nan X., Ng H.H., Johnson C.A., Laherty C.D., Turner B.M., Eisenman R.N., Bird A. (1998b). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393: 386-389
- Nan X., Tate P., Li E., Bird A. (1996). DNA methylation specifies chromosomal localization of MeCP2. *Molecular and Cellular Biology* 16: 414-421
- Neddermann P., Gallinari P., Lettieri T., Schmid D., Truong O., Hsuan J.J., Wiebauer K., Jiricny J. (1996). Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. *The Journal of Biological Chemistry* 271: 12767-12774
- Nelson H.C., Finch J.T., Luisi B.F., Klug A. (1987). The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature* 330: 221-226
- Ng J.D., Gavira J.A., Garcia-Ruiz J.M. (2003). Protein crystallization by capillary counterdiffusion for applied crystallographic structure determination. *Journal of structural biology* 142: 218-231
- Ohki I., Shimotake N., Fujita N., Jee J., Ikegami T., Nakao M., Shirakawa M. (2001). Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell* 105: 487-497
- Otani J., Arita K., Kato T., Kinoshita M., Kimura H., Suetake I., Tajima S., Ariyoshi M., Shirakawa M. (2013). Structural basis of the versatile DNA recognition ability of the methyl-CpG binding domain of methyl-CpG binding domain protein 4. *The Journal of Biological Chemistry* 288: 6351-6362
- Painter J., Merritt E.A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallographica Section D 62: 439-450
- Porath J., Carlsson J., Olsson I., Belfrage G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* 258: 598-599
- Potterton E., McNicholas S., Krissinel E., Cowtan K., Noble M. (2002). The CCP4 molecular-graphics project. *Acta Crystallographica. Section D, Biological Crystallography* 58: 1955-1957
- Prive G.G., Heinemann U., Chandrasegaran S., Kan L.S., Kopka M.L., Dickerson R.E. (1987). Helix geometry, hydration, and G.A mismatch in a B-DNA decamer. *Science (New York, N.Y.)* 238: 498-504
- Provencher S.W., Glockner J. (1981). Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* 20: 33-37

- Ramachandran G.N., Ramakrishnan C., Sasisekharan V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* 7: 95-99
- Raussens V., Ruysschaert J.M., Goormaghtigh E. (2003). Protein concentration is not an absolute prerequisite for the determination of secondary structure from circular dichroism spectra: a new scaling method. *Analytical Biochemistry* 319: 114-121
- Reese B.E., Bachman K.E., Baylin S.B., Rountree M.R. (2003). The methyl-CpG binding protein MBD1 interacts with the p150 subunit of chromatin assembly factor 1. *Molecular and Cellular Biology* 23: 3226-3236
- Reeves R., Nissen M.S. (1990). The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *The Journal of Biological Chemistry* 265: 8573-8582
- Robertson K.D. (2002). DNA methylation and chromatin unraveling the tangled web. *Oncogene* 21: 5361-5379
- Sansom O.J., Berger J., Bishop S.M., Hendrich B., Bird A., Clarke A.R. (2003). Deficiency of Mbd2 suppresses intestinal tumorigenesis. *Nature Genetics* 34: 145-147
- Sansom O.J., Maddison K., Clarke A.R. (2007). Mechanisms of disease: methyl-binding domain proteins as potential therapeutic targets in cancer. *Nature Clinical Practice. Oncology* 4: 305-315
- Scarsdale J.N., Webb H.D., Ginder G.D., Williams D.C., Jr. (2011). Solution structure and dynamic analysis of chicken MBD2 methyl binding domain bound to a target-methylated DNA sequence. *Nucleic Acids Research* 39: 6741-6752
- Sjolund A.B., Senejani A.G., Sweasy J.B. (2013). MBD4 and TDG: multifaceted DNA glycosylases with ever expanding biological roles. *Mutation Research* 743-744: 12-25
- Studier F.W., Moffatt B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology* 189: 113-130
- Toffaletti J., Savory J., Gitelman H.J. (1977). Use of gel filtration to examine the distribution of calcium among serum proteins. *Clinical chemistry* 23: 2306-2310
- Turkenburg J.P., Dodson E.J. (1996). Modern developments in molecular replacement. *Current Opinion in Structural Biology* 6: 604-610
- Vagin A., Teplyakov A. (2010). Molecular replacement with MOLREP. Acta crystallographica. Section D, Biological Crystallography 66: 22-25

- Vaguine A.A., Richelle J., Wodak S.J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallographica*. *Section D*, *Biological Crystallography* 55: 191-205
- Venkateswaran S., McMillan H.J., Doja A., Humphreys P. (2014). Adolescent onset cognitive regression and neuropsychiatric symptoms associated with the A140V MECP2 mutation. *Developmental Medicine and Child Neurology* 56: 91-94
- Wakefield R.I., Smith B.O., Nan X., Free A., Soteriou A., Uhrin D., Bird A.P., Barlow P.N. (1999). The solution structure of the domain from MeCP2 that binds to methylated DNA. *Journal of Molecular Biology* 291: 1055-1065
- Walavalkar N.M., Cramer J.M., Buchwald W.A., Scarsdale J.N., Williams D.C., Jr. (2014). Solution structure and intramolecular exchange of methyl-cytosine binding domain protein 4 (MBD4) on DNA suggests a mechanism to scan for mCpG/TpG mismatches. *Nucleic Acids Research*
- Wang S., Poon G.M., Wilson W.D. (2015). Quantitative investigation of protein-nucleic acid interactions by biosensor surface plasmon resonance. *Methods Mol Biol* 1334: 313-332
- Watanabe S., Ichimura T., Fujita N., Tsuruzoe S., Ohki I., Shirakawa M., Kawasuji M., Nakao M. (2003). Methylated DNA-binding domain 1 and methylpurine-DNA glycosylase link transcriptional repression and DNA repair in chromatin. Proceedings of the National Academy of Sciences of the United States of America 100: 12859-12864
- Weaving L.S., Christodoulou J., Williamson S.L., Friend K.L., McKenzie O.L., Archer H., Evans J., Clarke A., Pelka G.J., Tam P.P., Watson C., Lahooti H., Ellaway C.J., Bennetts B., Leonard H., Gecz J. (2004). Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *American Journal of Human Genetics* 75: 1079-1093
- Weiss M.S., Einspahr H., Baker E.N., Dauter Z., Kaysser-Pyzalla A.R., Kostorz G., Larsen S. (2010). Citations in supplementary material. Acta Crystallographica. Section D, Biological Crystallography 66: 1269-1270
- Wiebauer K., Jiricny J. (1989). In vitro correction of G.T mispairs to G.C pairs in nuclear extracts from human cells. *Nature* 339: 234-236
- Wilchek M., Bayer E.A. (1990). Introduction to avidin-biotin technology. *Methods in Enzymology* 184: 5-13
- Willard H.F., Hendrich B.D. (1999). Breaking the silence in Rett syndrome. *Nature Genetics* 23: 127-128
- Wilson G.G., Murray N.E. (1991). Restriction and modification systems. *Annual Review* of Genetics 25: 585-627

- Wing R., Drew H., Takano T., Broka C., Tanaka S., Itakura K., Dickerson R.E. (1980). Crystal structure analysis of a complete turn of B-DNA. *Nature* 287: 755-758
- Winn M.D., Ballard C.C., Cowtan K.D., Dodson E.J., Emsley P., Evans P.R., Keegan R.M., Krissinel E.B., Leslie A.G., McCoy A., McNicholas S.J., Murshudov G.N., Pannu N.S., Potterton E.A., Powell H.R., Read R.J., Vagin A., Wilson K.S. (2011). Overview of the CCP4 suite and current developments. Acta Crystallographica. Section D, Biological Crystallography 67: 235-242
- Wong E., Yang K., Kuraguchi M., Werling U., Avdievich E., Fan K., Fazzari M., Jin B., Brown A.M., Lipkin M., Edelmann W. (2002). Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation. Proceedings of the National Academy of Sciences of the United States of America 99: 14937-14942
- Wrinch D. (1946). Patterson distributions and native protein crystallography. *Nature* 157: 226
- Wu P., Qiu C., Sohail A., Zhang X., Bhagwat A.S., Cheng X. (2003). Mismatch repair in methylated DNA. Structure and activity of the mismatch-specific thymine glycosylase domain of methyl-CpG-binding protein MBD4. *The Journal of Biological Chemistry* 278: 5285-5291
- Xu X., Pozzo-Miller L. (2013). A novel DNA-binding feature of MeCP2 contributes to Rett syndrome. *Frontiers in cellular neuroscience* 7: 64
- Yesselman J.D., Horowitz S., Brooks C.L., 3rd, Trievel R.C. (2015). Frequent side chain methyl carbon-oxygen hydrogen bonding in proteins revealed by computational and stereochemical analysis of neutron structures. *Proteins* 83: 403-410
- Yusufzai T.M., Wolffe A.P. (2000). Functional consequences of Rett syndrome mutations on human MeCP2. *Nucleic Acids Research* 28: 4172-4179
- Zhao N., Ma D., Leong W.Y., Han J., Vandongen A., Chen T., Goh E.L. (2015). The methyl-CpG-binding domain (MBD) is crucial for MeCP2's dysfunctioninduced defects in adult newborn neurons. *Frontiers in Cellular Neuroscience* 9: 158