



**HAL**  
open science

# Elucidating the activation mechanism of the transcription factor DntR using X-ray crystallography and small angle X- ray scattering

Michael Lerche

► **To cite this version:**

Michael Lerche. Elucidating the activation mechanism of the transcription factor DntR using X-ray crystallography and small angle X- ray scattering. Structural Biology [q-bio.BM]. Université de Grenoble, 2014. English. NNT : 2014GRENV013 . tel-01555467

**HAL Id: tel-01555467**

**<https://theses.hal.science/tel-01555467>**

Submitted on 4 Jul 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

### DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 7 août 2006

Présentée par

**Michael LERCHE**

Thèse dirigée par **Gordon LEONARD**

codirigée par **Cyril DIAN**

préparée au sein du **European Synchrotron Radiation Facility**  
dans l'**École Doctorale Chimie et Science du Vivant**

**Elucidating the activation mechanism of  
the transcription factor DntR using X-ray  
crystallography and Small Angle X-ray  
Scattering**

**Compréhension du mécanisme  
d'activation du facteur de transcription  
DntR par cristallographie aux rayons X et  
diffusion des rayons X aux petits angles**

Thèse soutenue publiquement le **18 Julliet 2014**,  
devant le jury composé de :

**Mr. Serge PEREZ**

CERMAV Grenoble, Président

**Mme Nushin AGHAJARI**

IBCP Lyon, Rapporteur

**Mr Louis RENAULT**

CNRS Gif-Sur-Yvette, Rapporteur

**Mr Frank GABEL**

IBS Grenoble, Membre

**Mme Nicole THIELENS**

IBS Grenoble, Membre

**Mr Cyril DIAN**

IMAXIO Lyon, Membre

**Mr Gordon LEONARD**

ESRF Grenoble, Membre

*Université Joseph Fourier / Université Pierre Mendès France /  
Université Stendhal / Université de Savoie / Grenoble INP*



## Acknowledgements

First things first, a big thanks to Gordon and Cyril for letting me work on the DntR project. In particular a great thanks to Gordon for always being present with advice and encouragement while still allowing me to work as independently as I could have wanted. But mainly for always caring for both the project and me, it has really been a pleasure, thank you!

Thanks to everyone in the ESRF structural biology group and in particular to past and present members of the CIBB laboratory for providing a great working environment, which I was so happy to be a part of.

Thanks to the members of my thesis committee, in particular Chloe and Adam for willing or unwillingly accepting to be hassled with questions.

Thanks to Marie for joining in on the DntR project and being a pleasure to work with.

Tak til familie og venner i Danmark der paa trods af afstand tog sig til at besøge mig, forblev en del af mit liv igennem de sidste tre aar og altid var parate til at stille en seng, sofa eller bjernebryg til raadighed naar jeg kom forbi Danmark.

Thanks to the original lab core crowd of Erika, Tobias and Morgane who were always there for camping, movies, skiing, dinners, parties, hiking, etc, and for support in both the good times and the other ones.

Thanks to the later additions of Alex, Benoit, Taiana, Otilie, Gautier and Catarina for joining in.

Thanks to Gaëlle for daily coffees and running trips and for without whom this thesis would never have been completed.

Thanks to everyone who took part in making these last three years of my life unforgettable, I'll miss you all terribly.

## Table of Contents

Aknowledgements.....	1
Table of Contents .....	2
Popularized summary.....	5
Résumé de these vulgarise.....	5
Abstract .....	6
Résumé.....	8
Abbreviations .....	10
List of Figures.....	12
List of Tables.....	16
1. Introduction.....	17
1.1 Transcription .....	17
1.2 Transcription factors .....	18
1.3 LysR type transcription factors.....	21
1.3.1 Transcriptional regulation by LTTRs .....	23
1.3.2 The structure of LTTR proteins.....	25
1.3.3 The LTTR DNA Binding Domain. ....	29
1.3.4 LTTR Inducer Binding Domains.....	33
1.3.5 The crystal structure of full lengths LTTRs .....	37
1.3.6 Non tetrameric LTTRs.....	40
1.4 The LysR transcriptional regulator DntR .....	40
1.4.1 The structure of DntR.....	43
1.4.1.1 Inducer-bound DntR.....	46
1.5 The activation of LTTRs.....	47
1.6 Aim of thesis.....	49
2. Materials and methods .....	51
2.1 Expression and purification .....	51
2.1.1 Wild type DntR (DntR-His <sub>6</sub> ) .....	51
2.1.2 H169TDntR-His <sub>6</sub> .....	56
2.1.2.1 Mass spectrometry.....	58

2.1.3 $\Delta$ N90DntR and $\Delta$ N90H169TDntR.....	59
2.2 pDnT .....	60
2.3 Crystallization .....	63
2.3.1 DntR-His <sub>6</sub> .....	63
2.3.2 H169TDntR-His <sub>6</sub> .....	63
2.3.3 Salicylate-bound DntR-His <sub>6</sub> .....	63
2.3.3.1 High throughput crystallization trials.....	66
2.3.4 Crystal data collection and refinement.....	66
2.3.5 Thermofluor assays .....	68
2.4 Small Angle X-ray Scattering (SAXS).....	68
2.4.1.1 $\Delta$ N90DntR-His <sub>6</sub> and $\Delta$ N90H169TDntR-His <sub>6</sub> .....	68
2.4.1.2 DntR-His <sub>6</sub> .....	70
2.4.1.3 H169TDntR-His <sub>6</sub> .....	71
2.4.2 HPLC-coupled SAXS experiments .....	71
2.4.3 SAXS Data analysis.....	75
2.4.3.1 Ab initio modelling .....	75
2.4.3.2 Rigid body modelling.....	75
2.5 Figure preparation.....	77
3 Results .....	78
3.1 <i>apo</i> -DntR .....	78
3.1.1 The crystal structure of <i>apo</i> -DntR.....	78
3.1.2 Small Angle X-ray Scattering analysis <i>apo</i> -DntR .....	86
3.1.2.1 <i>ab initio</i> modelling.....	88
3.1.2.2 Rigid body modelling.....	91
3.2 H169TDntR .....	95
3.2.1 Crystallization of H169TDntR .....	95
3.2.2 Small Angle X-ray Scattering analysis H169TDntR .....	100
3.2.2.1 Molecular envelopes of the solution structures of <i>apo</i> -H169TDntR.....	102
3.2.2.2 Comparison of the results of SAXS experiments concerning <i>apo</i> -DntR and <i>apo</i> -H169TDntR .....	105
3.3 Conclusions on inducer-free DntR.....	105
3.4 Small Angle X-ray Scattering analysis of DntR in complex with salicylate .....	106

3.4.1	Solution structure of DntR pre-incubated with 100 $\mu$ M salicylate .....	110
3.4.2	Rigid body modeling.....	110
3.4.3	Conclusions.....	115
3.5	HPLC-coupled SAXS .....	115
4	Discussion .....	120
4.1	Summary of results .....	120
4.1.1	X-ray crystallography.....	120
4.1.1.1	<i>apo</i> -DntR .....	120
4.1.1.2	H169TDntR .....	120
4.1.2	SAXS.....	121
4.1.2.1	<i>apo</i> -DntR .....	121
4.1.2.2	H169TDntR .....	121
4.1.2.3	<i>holo</i> -DntR .....	122
4.1.2.4	Gel filtrated samples .....	122
4.2	The inactive and active states of DntR .....	123
4.3	The activation of LTTR proteins.....	128
4.4	Future work. ....	130
	Reference list.....	131
5.	Appendix: Small Angle X-Ray Scattering .....	139
5.1	Principles .....	139
5.1.1	Scattering intensities.....	140
5.2	$R_g$ and $I(0)$ .....	142
5.2.1	Considerations.....	143
5.3	Pair distribution function ( $P(r)$ ) .....	143
5.3.1	Considerations.....	144
5.4	<i>Ab initio</i> modeling .....	145
5.4.1	Considerations.....	147
5.5	CRY SOL fitting of experimental and theoretical scattering curves .....	148
5.6	Rigid body modeling.....	148

## **Popularized summary**

Proteins are an essential part of life and are responsible for a wide array of functions. The ability to produce the right protein in a given situation is a prerequisite for the survival of any organism. The production of proteins by organisms thus needs to be highly regulated.

In all organisms the initial step in protein production involves converting genomic information, contained within their DNA, into a blueprint (known as messenger RNA, mRNA) for a specific protein. This process is known as transcription and is regulated by a type of proteins called transcription factors. In order to understand the mechanism of this regulation it is essential to know the three dimensional structure of transcription factors.

This thesis work utilized the synchrotron based techniques X-ray crystallography and Small Angle X-ray Scattering to investigate the structure of the transcription factor DntR in order to elucidate how its activation regulates gene transcription and thus protein production.

## **Résumé de thèse vulgarisé**

Les protéines sont des éléments essentiels à la Vie et ont une large gamme de fonctions. La faculté de produire la bonne protéine dans la bonne situation est un pré-requis à la survie de tout organisme, et cette production nécessite une régulation stricte et coordonnée.

Dans tous les organismes, l'étape initiale de la synthèse de protéines implique de copier l'information génétique codée par un gène contenu dans l'ADN en la convertissant en ARN messenger. Ce procédé est appelé transcription et est régulé par différentes protéines qui comprennent les facteurs de transcription. La connaissance de la structure tridimensionnelle d'une protéine est souvent essentielle pour comprendre sa fonction, son mécanisme d'action et sa régulation à l'échelle moléculaire.

Ce travail de thèse a consisté à étudier la structure du facteur de transcription DntR et de comprendre comment son activation régule la transcription, en utilisant les techniques complémentaires de cristallographie aux rayons X et de diffusion de rayons X aux petits angles.

## Abstract

LysR type transcriptional regulatory (LTTR) proteins are the largest family of transcription factors amongst prokaryotes. In spite of the size of the family, structural information on full-length constructs of these proteins is very limited as they are often insoluble and very difficult to crystallize. From the few existing crystal structures, coupled with other biophysical evidence, it is known that the proteins mainly associate as homotetramers comprising a dimer of dimers. The dimers associate through large C-terminal domains in a “head-to-tail” fashion and are connected “head-to-head” through their N-terminal domains and the resulting homotetramers are activated by the binding of inducer molecules. Each C-terminal domain contains an inducer binding cavity (IBC) and is denoted an inducer binding domain (IBD), while the N-terminal dimers each bind a region of DNA via a winged helix-turn-helix (wHTH) motif.

Unlike other transcription factors, LTTR proteins do not regulate expression by associating or disassociating with DNA. They bind to DNA in both their active and inactive states and the current consensus is that they regulate gene expression through large conformational changes that relax the bending of bound DNA. However, to this date, no crystal structures of a full length homotetrameric LTTR in both an active and inactive conformation exist, and thus their mechanism of transcriptional regulation remains structurally uncharacterized.

The work described in this thesis has used the LTTR DntR as a model protein to further structurally characterize the activation mechanism of LTTR proteins. The first crystal structure of *apo*-DntR is presented as is the crystal structure of H169TDntR, a mutant which shows activity in the absence of an inducer molecule. Thermofluor assays performed on this mutant, show that it has a melting temperature similar to that of inducer bound DntR. Comparison of these crystal structures with the crystal structure of salicylate-bound DntR reveals that the protein in its *apo*-state adopts a compact IBC, which precludes the binding of an inducer molecule. Despite the evidence of thermofluor assays, the crystal structure of H169TDntR is very similar to that of *apo*-DntR suggesting that crystal packing effects impose strong limitations on the use of crystallography to elucidate the active and inactive conformations of DntR. Small Angle X-ray Scattering (SAXS) was thus used to study the structure of DntR in solution.

SAXS study reveals that in solution DntR in its inactive *apo*-state is found in a different conformation compared to that seen in its crystal structure. While maintaining a compact tetrameric C-terminal core the DNA binding wHTH dimers pack much closer to this than seen in the crystal structure and adopt a conformation that would result in much higher bending of bound DNA than previously postulated.

SAXS studies of the constitutively active H169TDntR mutant confirm, as thermofluor assays had suggested, that in solution the structure of this protein is markedly different from its crystal structure. Indeed the solution structure of H169TDntR appears very like that of open-form homotetramers seen in the crystal structure of Tsar. This same effect was observed in solution scattering studies of inducer bound-and thus activated, DntR.

The work presented in this thesis thus appears to confirm, as previously hypothesized, that upon activation DntR, and presumably all homotetrameric LTTRs, undergo a conformational change from a compact, to a much more open form that allows the relaxation of the bound DNA promoter region, exposing it to solvent and allows RNA polymerase access and thus initiate transcription.

## Résumé

Les protéines régulatrices de la transcription de type LysR (LTTR) appartiennent à la plus grande famille de facteur de transcription (FT) chez les procaryotes. Malgré l'importance de cette famille, il existe peu de données structurales sur les LTTRs car ces FTs sont souvent insolubles et difficiles à cristalliser. Les quelques structures tridimensionnelles existantes et les différentes études biochimiques menées jusqu'à présent ont montré que ces FTs s'associent principalement sous forme d'homotétramère comprenant un dimère de dimères. Chaque dimère s'assemble en « tête-bêche » par un large domaine C-terminal. Ces dimères s'associent à leur tour par leurs domaines N-terminaux formant deux dimères en « tête-à-tête ». L'homotétramère ainsi formé est activé par la liaison de molécules effectrices dans une cavité (Inducer Binding Cavity : IBC) localisée dans le domaine C-terminal, nommé domaine de liaison à l'inducteur (Inducer Binding Domain : IBD). Chacun des dimères formés par les domaines N-terminaux se lie à une région de l'ADN par leurs motifs Hélice-Tour-Hélice ailés (winged HTH).

Contrairement aux autres facteurs de transcription, les LTTRs ne régulent pas l'expression protéique par une simple association ou dissociation avec l'ADN. Ces FTs sont continuellement liés à l'ADN dans leur état actif et inactif. Il existe un consensus sur la manière dont les LTTRs régulent la transcription. Cette régulation repose sur d'importants changements conformationnels qui induisent le relâchement de la courbure de l'ADN en amont du promoteur du gène régulé. À ce jour, aucun membre de la famille des LTTRs n'a vu ses deux conformations, active et inactive, déterminées par cristallographie aux rayons X. Par conséquent, leur mécanisme moléculaire de régulation transcriptionnelle reste à caractériser.

Le travail décrit dans cette thèse concerne un membre de la famille des LTTRs, DntR. La structure cristalline de DntR sous forme *apo* ou inactive est présentée ainsi que la structure du mutant H169T-DntR, qui est constamment actif en absence de molécules effectrices. L'analyse par fluorimétrie différentielle à balayage (DSF ou TSA) montre que la température de dénaturation de ce mutant autoinducteur est similaire à celle obtenue avec DntR liée à une molécule inductrice, le Salicylate. La comparaison des structures obtenues avec celles de l'IBD de DntR lié au Salicylate révèle que la protéine dans son état *apo* arbore une IBC

reserrée, qui ne favorise pas la liaison de molécules inductrices. Malgré l'évidence des informations apportées par le TSA, la structure du mutant autoinducteur H169T-DntR adopte, elle aussi, une structure quaternaire identique à DntR sous forme *apo*. La nécessité de la création de contacts lors de l'empilement cristallin impose d'importantes limitations à l'utilisation de la cristallographie aux rayons X pour clarifier les conformations active et inactive de DntR. Pour se libérer de ces restrictions, nous avons étudié la structure de DntR en solution par méthode de diffusion des rayons X aux petits angles (SAXS).

L'étude SAXS effectuée sur DntR montre que dans son état inactif, la conformation *apo* adopte une conformation compacte légèrement différente de celle observée dans la structure cristalline. Dans cette structure en solution, les dimères de wHTHs sont accolés au corps compact formé par les domaines C-terminaux et leurs orientations s'en trouvent modifiées. Le rapprochement des domaines de liaisons à l'ADN suggère que la courbure de l'ADN induite par DntR dans cette conformation serait plus importante que celle postulée précédemment.

Les études SAXS menées sur le mutant autoinducteur H169TDntR confirme l'analyse effectuée par TSA et montre que la structure quaternaire de ce mutant en solution est différente de celle obtenue dans le cristal. En effet, la structure en solution de H169TDntR est très similaire à la conformation ouverte de l'homotétramère observée dans la structure cristalline de Tsar. Cette conformation ouverte est aussi observée par méthode de diffusion des rayons X aux petits angles pour DntR activée en présence de molécules inductrices.

Les études SAXS sur DntR sauvage et son mutant autoinducteur H169TDntR en solution confirment l'hypothèse du mécanisme d'activation postulée pour la famille des LTTRs. Lors de la liaison d'une molécule effectrice, l'homotétramère subirait un important changement de conformation traduit par le repositionnement des domaines de liaison à l'ADN d'une forme compacte vers une forme ouverte induisant le relâchement de la courbure de la molécule d'ADN liée. Ce relâchement favoriserait l'accès de l'ARN polymérase aux régions promotrices et ainsi, l'initiation de la transcription de l'opéron régulé.

## Abbreviations

A <sub>260</sub>	Absorption at $\lambda=260$ nm
A <sub>280</sub>	Absorption at $\lambda=280$ nm
ABS	Activating binding site
AMP	Adenosine mono-phosphate
bp	Base pair
C-terminal	Carboxyl terminal
DBD	DNA binding domain
DNA	2'-Deoxyribonucleic acid
DNT	2,4-Dinitrotoluene
DTT	Dithiothreitol
$\epsilon$	Molar extinction coefficient
<i>E. coli</i>	Escherichia Coli
EDTA	Ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
ESRF	European Synchrotron Radiation Facility
GFP	Green fluorescent protein
His <sub>6</sub>	Hexa-histidine tag
HTH	Helix-turn-helix domain
HTX	High throughput crystallization laboratory
HPLC	High pressure liquid chromatography
IBC	Inducer binding cavity
IBD	Inducer binding domain
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
ITC	Isothermal calorimetry

LB	Luria broth medium
LTRR	LysR-type transcriptional regulator
N-terminal	Amino terminal
Ni-NTA	Nickel-nitriloacetic acid
OD <sub>600</sub>	Optical density at $\lambda=600$ nm
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
RBS	Regulatory/Repressing binding site
RD	Rossmann-fold like domain
RMSD	Root mean square deviation
RNA	Ribonucleic acid
RNA pol	RNA polymerase
SAXS	Small Angle X-ray Scattering
SANS	Small Angle Neutron Scattering
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
TBE	Tris-Boric Acid-EDTA buffer
TCA	<i>p</i> -toluenecarboxylate
TF	Transcription factor
TRIS	2-amino-2-hydroxymethyl-propane-1,3-diol
TSA	<i>p</i> -toluenesulfonate
v/v	volume per volume
wHTH	Winged helix-turn-helix domain
Å	Ångstrom (1 Å = 0.1 nm)

## List of Figures

- Figure 1.1: The crystal structure of RNA polymerase-II
- Figure 1.2: The crystal structure of the *Escherichia Coli* catabolite activator protein (CAP), PDB: 3GAP
- Figure 1.3: Schematics showing one and two component regulatory systems
- Figure 1.4: Examples of positive and negative transcriptional regulation
- Figure 1.5: Examples of gene regulation via TF concentration
- Figure 1.6: The diversity of gene products regulated by LTTR proteins in *S. Enterica*
- Figure 1.7: The classical model of LTTR dependent transcriptional regulation
- Figure 1.8: A schematic showing the current model of LTTR transcriptional regulation
- Figure 1.9: The domain structure of LTTR monomers
- Figure 1.10: Proposed oligomerisation schemes for the LTTR BenM
- Figure 1.11: The pathway for the degradation of benzoate and anthranilate in the soil bacterium *Acinetobacter bayly*
- Figure 1.12: The pathway for the degradation of TSA and TCA in *Comamonas Testosterone*
- Figure 1.13: The crystal structure of dimeric DBD/linker helix domains of LTTR homotetramers
- Figure 1.14: Amino acid sequence alignment of the DBD/linker helix domains of various LTTRs
- Figure 1.15: The binding of a dimer of the BenM DBD linker helix domain to its promoter region DNA
- Figure 1.16: Amino acid sequence alignment of the IBDs of selected LTTRs
- Figure 1.17: The crystal structures of the LTTR IBD monomer
- Figure 1.18: The crystal structures of LTTR IBD dimers
- Figure 1.19: The crystal structure of a BenM monomer with both benzoate and muconate-bound
- Figure 1.20: Schematic of the crystal structure of a typical LTTR homotetramer
- Figure 1.21: Crystal structures of open and compact full length LTTR homotetramers
- Figure 1.22: The proposed oligomerisation and DNA binding scheme of CrgA
- Figure 1.23: DNT degradation in *Burkholderia sp.*
- Figure 1.24: The arrangement of the catabolic genes in the DNT degrading pathway in *Burkholderia sp.*
- Figure 1.25: The premise of a biosensor for DNT
- Figure 1.26: The domain structure of DntR
- Figure 1.27: The current available crystal structure of the full length DntR homotetramer
- Figure 1.28: Comparison of the crystal structures of DntR and CbnR homotetramers
- Figure 1.29: The crystal structure of salicylate-bound  $\Delta$ N90-DntR (PDB 2Y7K)
- Figure 1.30: Conformational changes to the DntR IBD upon inducer binding
- Figure 2.1: The pQE-60 plasmid
- Figure 2.2: Purification of DntR-His<sub>6</sub> by nickel affinity chromatography

Figure 2.3: Purification of DntR-His<sub>6</sub> by gel filtration

Figure 2.4: SDS PAGE analysis of purified DntR-His<sub>6</sub>

Figure 2.5: Size exclusion chromatography elution profiles of DntR-His<sub>6</sub>

Figure 2.6: Purification of H169TDntR-His<sub>6</sub> by nickel affinity chromatography

Figure 2.7: Deconvoluted MALDI-TOF spectrum of H169TDntR-His<sub>6</sub>.

Figure 2.8: Purification of ΔN90DntR-His<sub>6</sub> by nickel affinity chromatography

Figure: 2.9: The PCR reaction for the creation of the pDnT fragment

Figure 2.10: PCR reactions and pDnT DntR binding

Figure 2.11: Crystals of full-length DntR grown in the presence of salicylate

Figure 2.12: Cumulative Intensity Distribution plots for crystals of *apo*-DntR-His<sub>6</sub> and DntR-His<sub>6</sub> grown in the presence of salicylate

Figure 2.13: Crystals of potential salicylate-bound DntR-His<sub>6</sub> obtained from high throughput crystallization screening

Figure 2.14: The crystallographic data collection, analysis and refinement strategy used.

Figure 2.15.: Scattering curves for ΔN90DntR-His<sub>6</sub> and ΔN90H169TDntR-His<sub>6</sub>

Figure 2.16: Scattering curves from solutions of ΔN90 DntR-His<sub>6</sub> construct in the absence and presence of 5 mM salicylate

Figure 2.17: *ab initio* model of the solution structures of ΔN90DntR-His<sub>6</sub>

Figure 2.18: The setup for HPLC-coupled BioSAXS at ESRF beamline BM29

Figure 2.19: Summarized output file for HPLC-coupled SAXS experiments of solutions of *apo*-DntR-His<sub>6</sub>

Figure 2.20: Summarized output file for HPLC-coupled SAXS experiments of solutions of *holo*-DntR-His<sub>6</sub>

Figure 2.21: Flowchart of SAXS data processing

Figure 2.22: Rigid body modelling of the solution structures of *apo*- and *holo*-DntR-His<sub>6</sub>

Figure 3.1: Electron density for the N-terminal regions in the crystal structure of *apo*-DntR

Figure 3.2: The crystal structure of *apo*-DntR.

Figure 3.3: A superposition of the crystal structures of homotetrameric *apo*-DntR and thiocyanate-bound DntR

Figure 3.4: Electron density for the IBC in the crystal structure of *apo*-DntR

Figure 3.5: A superposition of IBCs of the crystal structures of tetrameric *apo*-DntR and of acetate-/thiocyanate-bound DntR

Figure 3.6: Comparison of the conformations of the IBCs of *apo*-DntR and *holo*-ΔN90-DntR

Figure 3.7: The solution scattering curve and corresponding Guinier region for *apo*-DntR

Figure 3.8: The P(r) function of the solution structure of *apo*-DntR

Figure 3.9: SAXS analysis of *apo*-DntR - CRY SOL fits of experimental and theoretical scattering curves

Figure 3.10: SAXS-derived *ab initio* molecular envelopes for the solution structure of the *apo*-DntR homotetramer overlaid with the crystal structure of the *apo*-DntR homotetramer

Figure 3.11: The rigid body composition of the crystal structure of DntR used in SASREF

Figure 3.13: SASREF-derived solution structure and *ab initio* envelopes for homotetrameric *apo*-DntR

Figure 3.14: Orthogonal views of solution and crystal structures of *apo*-DntR.

Figure 3.15: Thermofluor assays of *apo*- and *holo*-DntR and *apo*-H169TDntR

Figure 3.16: Crystals of H169TDntR

Figure 3.17: The crystal structure of H169TDntR

Figure 3.18: Superposition of the IBCs seen in the crystal structures of *apo*-DntR, H169TDntR and *holo*- $\Delta$ N90DntR

Figure 3.19: SAXS scattering curve and corresponding Guinier region and P(r) function for *apo*-H169TDntR

Figure 3.20: SAXS analysis of *apo*-H169TDntR - CRY SOL fits against the experimental scattering data.

Figure 3.21: *ab initio* envelope of the solution structure of H169TDntR

Figure 3.22: Solution scattering curves obtained from solutions of DntR pre-incubated with various concentrations of sodium salicylate

Figure 3.23: The Guinier regions of scattering curves obtained from solutions of DntR pre-incubated with salicylate

Figure 3.24: P(r) functions of the scattering curves shown in Figure 3.21

Figure 3.25: SAXS determined *ab initio* molecular envelope of salicylate-bound DntR overlaid with different models of the structure of full length DntR tetramer

Figure 3.26: SAXS analysis of a solution of DntR pre-incubated with 100  $\mu$ M sodium salicylate - CRY SOL fits against the experimental data

Figure 3.27: The SASREF-obtained solution structure of salicylate-bound DntR

Figure 3.28: SAXS analysis of salicylate-bound DntR - CRY SOL fits against the experimental data

Figure 3.29: The online gel filtration elution profiles of solutions of *apo*- and *holo*-DntR

Figure 3.30: Averaged solution scattering curves for gel filtrated *apo*-(top) and *holo*-DntR (bottom)

Figure 3.31: Guinier plots and P(r) functions obtained of *apo*-DntR and DntR pre-incubated with 5 mM salicylate

Figure 4.1: Orthogonal views of the solution structures of *apo*- and *holo*-DntR. SASREF-generated solution structures of *apo*-DntR(left) and *holo*-DntR (right) homotetramers

Figure 4.2: Side views of the models of DNA binding position of the N-terminal WHTH/linker helix domains in the various DntR homotetramers observed

Figure 4.3: The possible displacement of a binding site for bound DNA upon DntR activation

Figure 4.4: Comparison the solution structure of *holo*-DntR the crystal structure of and Tsar

Figure 4.7: LTTR activation as given in the main text.

Figure 5.1: Setup of a SAXS experiment

Figure 5.2: Measuring SAXS scattering curves

Figure 5.3: Attractive and repulsive interactions

Figure 5.4: AUTOGNOM Guinier region estimation in primusqt

Figure 5.5: Scattering intensities and pair distribution functions of common geometrical bodies with the same  $D_{\max}$  values

Figure 5.6: *ab initio* modelling in the program DAMMIN

Figure 5.7: Examples of *ab initio* generated models

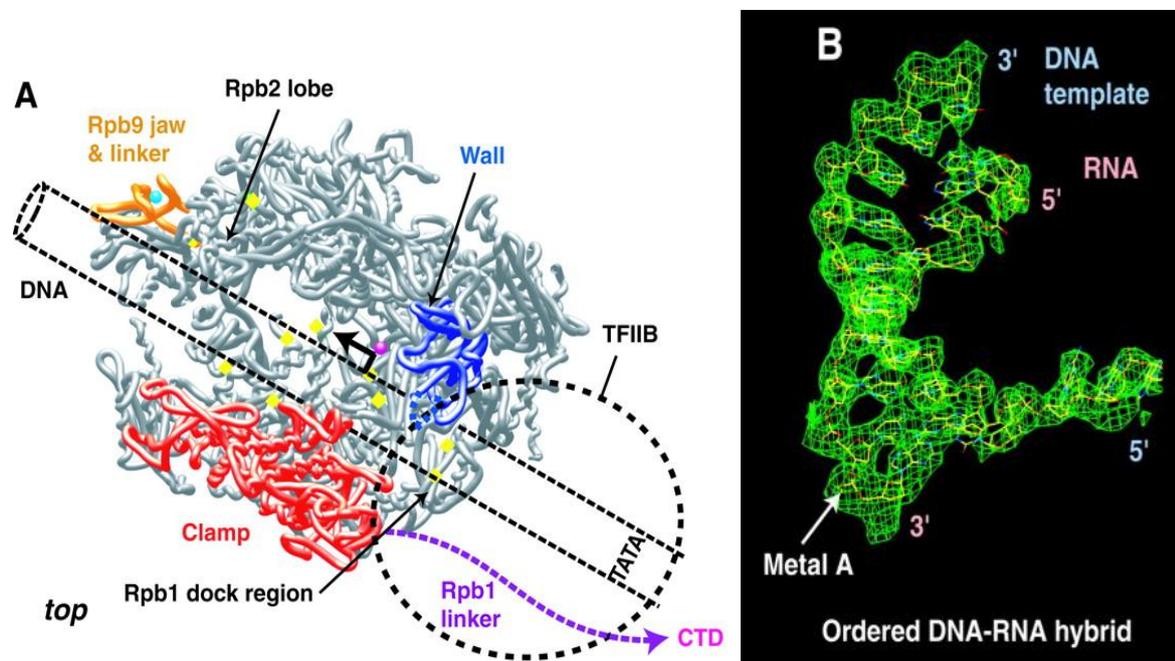
## List of Tables

- Table 1.1: Sequence identity matrix of LTTRs for which the crystal structures have been solved
- Table 1.2: Available crystal structures of LTTR proteins
- Table 1.3: Sequence identity matrix of the N-terminal of LTTR proteins
- Table 1.4: Sequence identity matrix of the C-terminal of LTTR proteins
- Table 2.1: Combinations of induction temperatures and hours of expression trialled
- Table 2.2: The PCR procedure used to create H169TDntR.
- Table 2.3: The PCR procedure used to create  $\Delta$ N90H169TDntR.
- Table 2.4: The PCR reactions used to create the pDnT fragment.
- Table 2.5: Data collection and refinement statistics for crystals of DntR-His<sub>6</sub> grown in the presence of salicylate
- Table 2.6: Invariant parameters obtained from the scattering curves of  $\Delta$ N90DntR constructs.
- Table 3.1: Data collection and refinement statistics for the crystal structure of *apo*-DntR
- Table 3.2:  $\chi^2$  values of the CRY SOL fits shown in Figure 3.9 and 3.12.
- Table 3.3: Data collection and refinement statistics for the crystal structure of H169TDntR
- Table 3.4: Comparison of invariant parameters obtained from solution scattering curves for *apo*-DntR and *apo*-H169TDntR.
- Table 3.5:  $\chi^2$  values obtained for the CRY SOL fits shown in Figure 3.21.
- Table 3.6: Invariant parameters obtained from SAXS scattering curves from solutions of *apo*-DntR and of DntR pre-incubated with different concentrations of sodium salicylate
- Table 3.7: Invariant parameters obtained from gel filtrated SAXS scattering curves of DntR in the absence and presence of 5 mM sodium salicylate.
- Table 4.1: invariant SAXS parameters obtained for *apo*- and *holo*-DntR with and without gel filtration prior to measurements.

# 1. Introduction

## 1.1 Transcription

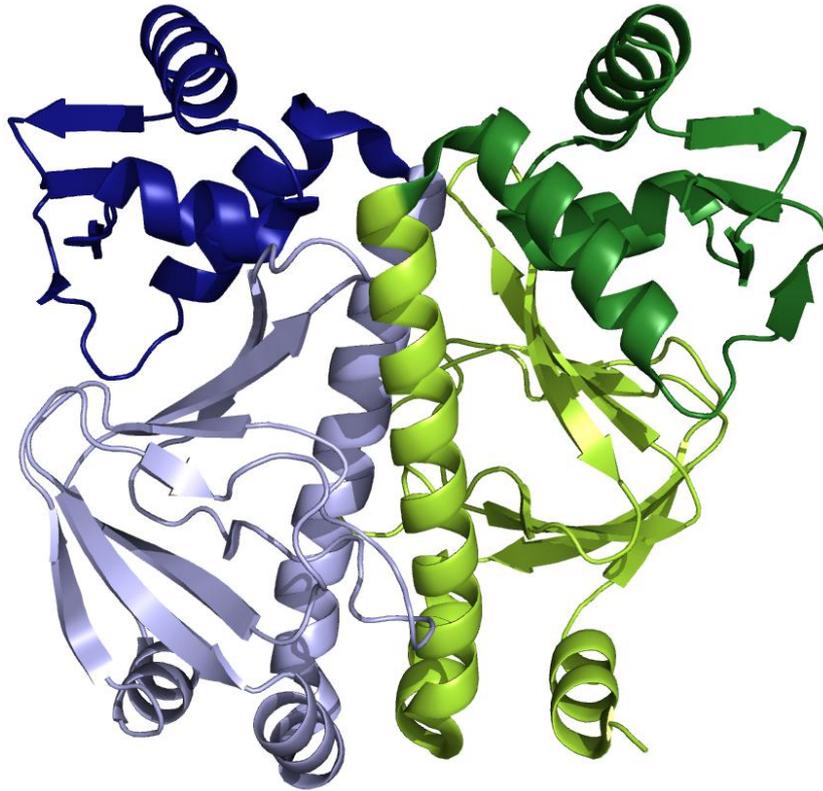
Transcription is the first step in gene expression in which a segment of DNA is copied into an mRNA transcript that is then decoded by ribosomes to produce proteins. The principal enzyme involved in the process of transcription is RNA polymerase, which catalyzes the polymerization of ribonucleoside 5'-triphosphates in the 5' to 3' direction. RNA polymerase initially binds to a promoter region located upstream of the gene to be transcribed (Cramer et al., 2001; Gnatt et al., 2001). The crystal structure of RNA Polymerase II from *Saccharomyces cerevisiae* is shown in Figure 1.1.



**Figure 1.1: The crystal structure of RNA polymerase-II.** A: The crystal structure of RNA polymerase II with the proposed position of bound DNA. As published by (Cramer et al., 2001). B: Electron density for a DNA-RNA hybrid bound to RNA-polymerase. The location of the active site metal A is indicated. As published by (Gnatt et al., 2001).

A major strength of prokaryotes is their ability to regulate and fine tune gene expression to respond to internal and external stimuli. The fact that gene expression was regulated was discovered in the late 1950s during studies of the induction of the *lac* operon (Jacob and Monod, 1959) and the control of the lytic-lysogenic cycle in  $\lambda$ -phage infection (Ptashne, 1965). Multiple subsequent studies showed that there were generalities as to how cells coupled various stimuli with gene expression (Burgess et al., 1969; Stevens and Rhoton,

1975), but mechanisms via which gene expression could be regulated, were not elucidated until the first crystallographic structures of so-called transcription factors (proteins that regulate gene expression) became available (McKay and Steitz, 1981; Weber and Steitz, 1987) (Figure 1.2).

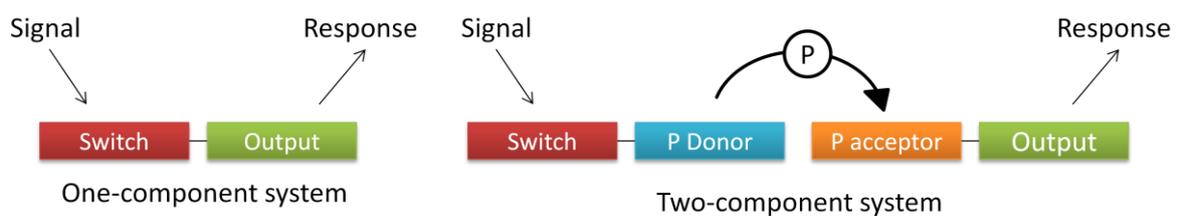


**Figure 1.2:** The crystal structure of the *Escherichia Coli* catabolite activator protein (CAP), PDB: 3GAP. The protein associates as a dimer (monomers shown in green and blue) with cyclic AMP binding sites located in the C-terminal switch domain (light blue and light green) while the N-termini (dark blue and dark green) were presumed to bind DNA (Weber and Steitz, 1987).

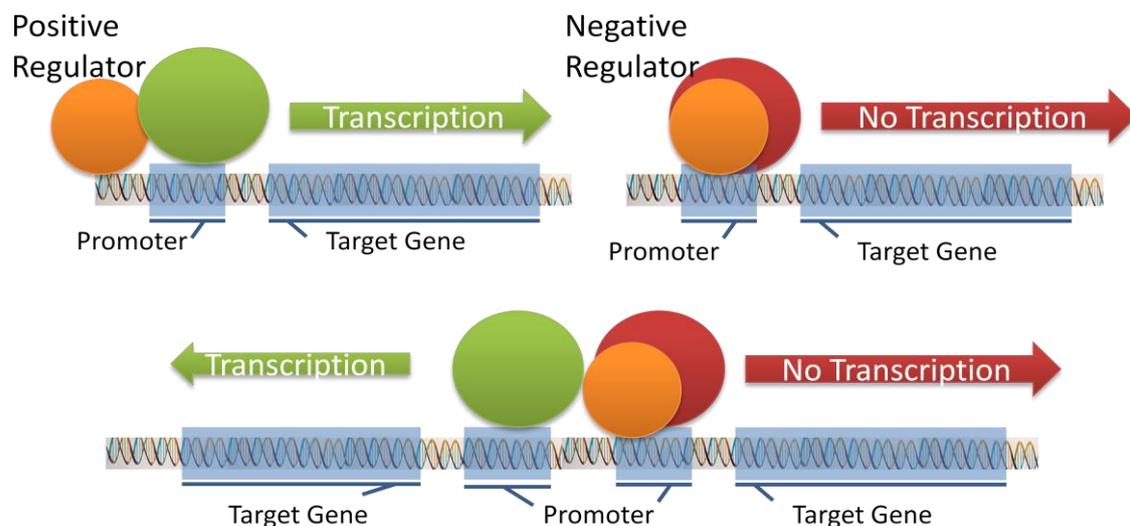
## 1.2 Transcription factors

As shown in Figure 1.2, transcription factors (TFs) require at least two domains to allow them to function as regulatory proteins (Balleza et al., 2009). One domain serves as switch, responding to ligand binding, protein-protein interactions or physiochemical signals such as a change in pH or temperature (Martínez-Antonio et al., 2006; Ptashne and Gann, 2001). The second domain directly binds a target DNA sequence close to a promoter region and will respond according to whether the switch domain is activated/inactivated (Balleza et al., 2009). In bacteria the DNA binding domain usually consists of the helix-turn-helix (HTH)

motif which has been identified in 95% of all prokaryotic DNA binding proteins (Huffman and Brennan, 2002). The two domains required by a TF can be contained either within a single protein (one-component system) or two different proteins (two-component system) (Ulrich et al., 2005). In the latter the switch domain is usually located in a membrane-linked multidomain sensor kinase which can then communicate, through phosphorylation cascades, with its cytosolic DNA binding partner protein (Mascher, 2006). The two types of TFs are schematized in Figure 1.3.



**Figure 1.3: Schematics showing one and two component regulatory systems.** A one-component system (left) is a single protein that contains both switch (red) and output (green) domains. Two-component signal transduction systems contain their switch and output domains in different proteins which communicate via phosphotransfer domains (blue and orange). Figure from adapted from (Ulrich et al., 2005).

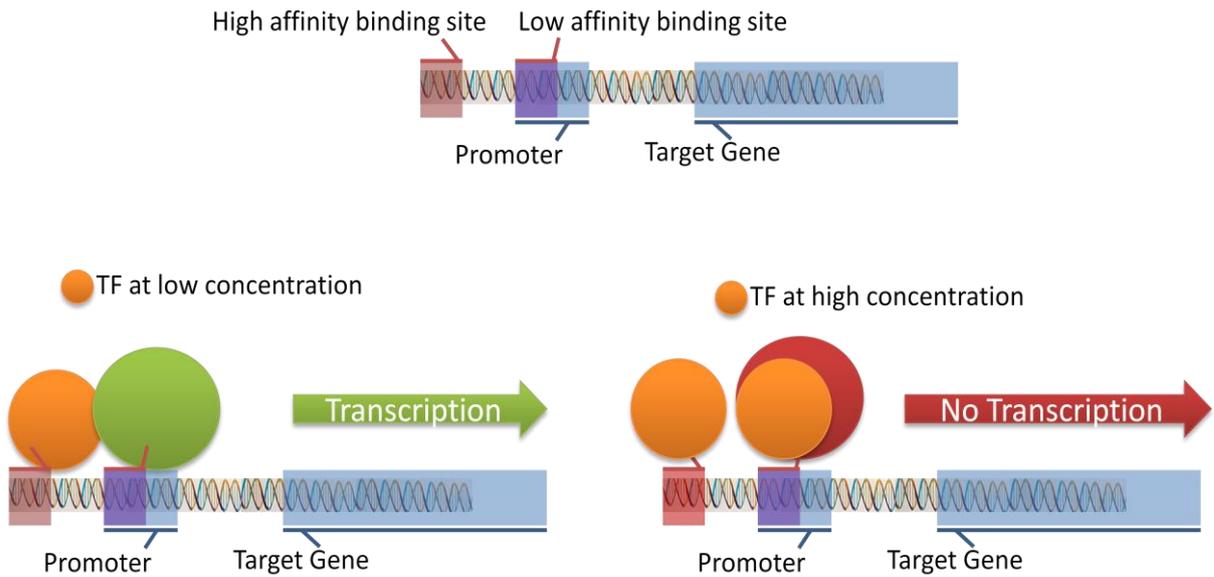


**Figure 1.4: Examples of positive and negative transcriptional regulation.** Top left: A positive regulator of transcription (orange) binds upstream of the promoter region and helps to recruit RNA polymerase (green). Top Right: A negative regulator (orange) binds on the promoter region and blocks RNA polymerase (red) from binding thus preventing transcription. Bottom: A transcription factor (orange) binds in the intergenic region between divergently transcribed genes and recruits RNA polymerase (green) allowing for the transcription of one gene, while RNA polymerase binding to the second promoter region, and thus transcription of the second gene, is blocked.

TFs can be divided into two further subcategories: negative and positive regulators (Figure 1.4, top). A negative regulator will usually bind DNA at, or very near to, a promoter region and thereby sterically prevents RNA polymerase from accessing the DNA (Collado-Vides et al., 1991). Positive regulators bind upstream of the promoter region and help to recruit the polymerase (Madan Babu and Teichmann, 2003). Some transcription factors may perform both roles, binding to the intergenic region between divergently transcribed genes, affecting regulation of both of them in different manners (Figure 1.4, bottom) (Balleza et al., 2009).

Bacterial TFs usually operate as homomultimers of various orders and only rarely as heteromultimers (Goulian, 2004). TFs are promiscuous to various degrees and can be classified as either Local or Global TFs: Local TFs tend to bind only to one specific DNA sequence while Global TFs can bind to a larger ensemble of DNA regions (Lozada-Chávez et al., 2008; Martínez-Antonio et al., 2006) and thus be involved in the regulation of more than one operon. Gene expression is also affected by the relationship between the affinity of the TF to the DNA binding site and the relative expression the TF within the cell. In fact, a low affinity TF may, at higher concentrations, regulate expression of genes normally regulated by higher affinity TFs. TFs can also perform a dual role based on concentration in the cell (Figure 1.5). If a gene contains a binding site inside the promoter with a weak binding affinity and a binding site upstream of the promoter of strong binding affinity, at low concentrations the TF will only bind the high affinity site and thus promote transcription, at higher concentrations the low affinity binding site will also be occupied and transcription will thus be blocked (Balleza et al., 2009).

A strong correlation has been established between the relative stability of the ecological niche that a bacterium occupies and the proportion of transcriptional regulators, specifically TFs, encoded in a genome. Bacteria that live in relatively stable niches, such as intracellular pathogens, endosymbionts and extremophiles, have undergone a reduction in genome size and harbour less TFs per gene. In contrast, bacteria that are exposed to extreme variation in their environment, such as free-living organisms and bacteria that colonise plants or infect animal tissues, possess large genomes with a larger TF-to-gene proportion. The evolutionary response produced by prokaryotes to a changing environment thus appears to rely to a large extent on transcriptional control of gene expression by TFs (Cases and de Lorenzo, 2005; Cases et al., 2003).

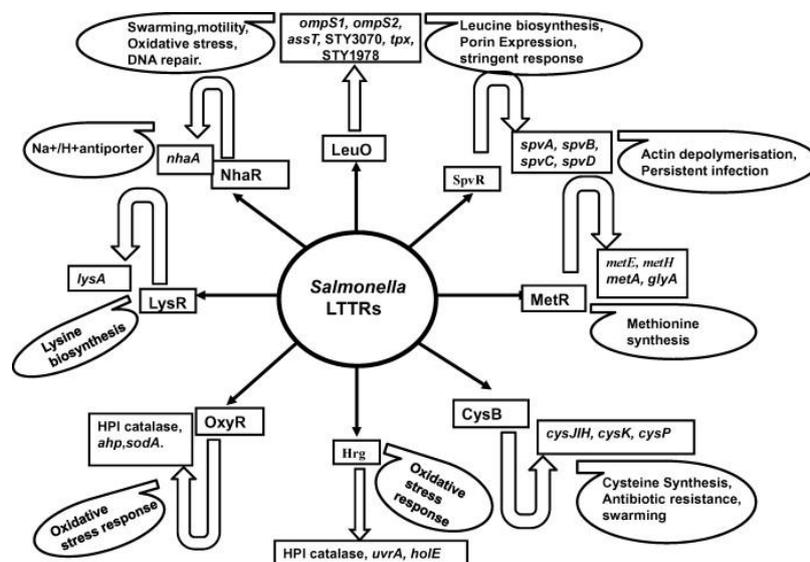


**Figure 1.5: Examples of gene regulation via TF concentration.** Top: A schematic of a gene containing two binding sites for a TF. One is a high affinity site (red) located upstream of the promoter, the other a low affinity site (purple) located within the promoter region. When the TF is present within the cell at low concentrations only the high affinity site will be occupied (bottom left) and the TF will act as a positive regulator. At higher concentrations the TF will also bind to the low affinity site (bottom right) and act as a negative regulator.

### 1.3 LysR type transcription factors

LysR-type transcriptional regulators (LTTR) were initially documented in 1988 by Henikoff et al (Henikoff et al., 1988) who, in various organisms (*Escherichia coli*, *Salmonella enterica*, *Rhizobium* and *Enterobacter cloacae*), identified regulatory proteins which, based on sequence homology and the predicted secondary structure of their DNA binding domains, could be denoted as a distinct class of bacterial transcriptional regulators. The group is named after one of its earliest identified members, LysR, which controls the transcription of the *lysA* operon that codes for enzymes which catalyze the decarboxylation of diaminopimelate to produce lysine (Stragier and Patte, 1983). LTTRs are present in a wide and diverse range of prokaryotes, suggesting their progenitor arose early in prokaryote evolution (Schell, 1993). Based on genome sequences more than 800 LTTRs have now been identified and the group comprises the largest known family of prokaryotic transcriptional regulators, corresponding to about 20% of all known bacterial TFs (Pareja et al., 2006).

The genes regulated by LTTRs usually express enzymes involved in basic metabolic pathways, but may also play a role in several other functions such as cell division, motility and various stress responses (reviewed in Maddocks and Oyston, 2008). By way of an example, the diversity of gene regulation by LTTRs in *S. enterica* is illustrated in Figure 1.6. Despite the size of the LTTR family, the variety of functions of the gene products of which they regulate expression, and a relatively low sequence identity (Table 1.1), the family shares a high degree of structural identity (Maddocks and Oyston, 2008) and LTTRs are thus expected to regulate transcription in a similar manner.



**Figure 1.6: The diversity of gene products regulated by LTTR proteins in *S. Enterica*.** Of the 49 reported LTTR proteins in *S. enterica*, the roles of 24 have been characterized. Figure as published by (Lahiri et al., 2009).

	BenM	CatM	CbnR	CynR	OxyR	TsaR	CysB	PA01	AphB	RovM	MetR
BenM											
CatM	58.75										
CbnR	29.35	27.99									
CynR	24.16	24.16	25.09								
OxyR	21.00	20.40	20.27	24.08							
TsaR	20.55	19.52	20.42	23.21	20.95						
CysB	19.00	19.40	17.87	23.08	19.54	20.61					
PA01	26.17	26.17	22.41	20.13	19.67	18.71	22.26				
AphB	20.21	19.16	18.25	19.51	13.45	18.34	14.88	19.45			
RovM	18.66	19.79	20.07	22.78	18.28	19.06	19.59	18.82	17.02		
MetR	21.40	21.48	17.59	19.13	18.36	17.69	19.02	22.52	15.07	17.87	
DntR	15.07	15.07	14.19	17.18	15.70	17.53	15.02	14.81	11.90	18.02	17.23

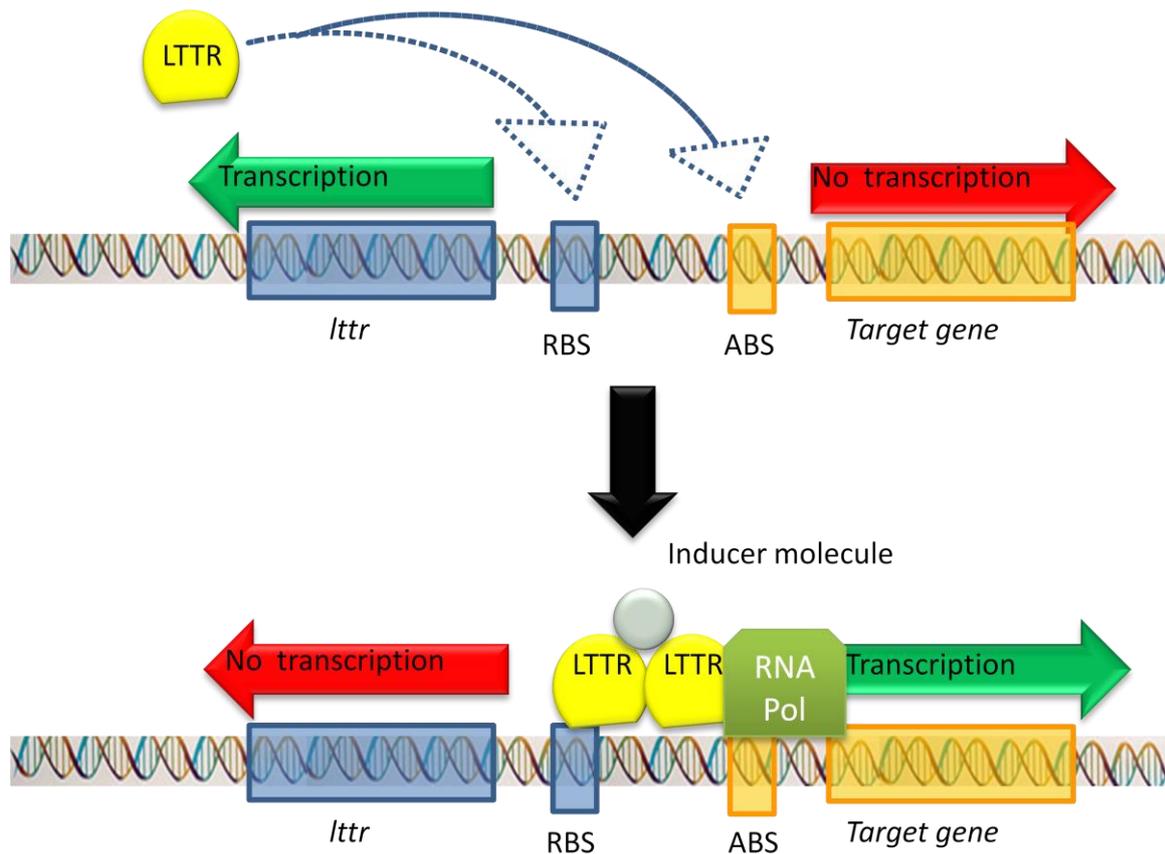
**Table 1.1: Sequence identity matrix of LTTRs for which the crystal structures have been solved.** The alignment was performed using the EMBL Clustalw2 server <http://www.ebi.ac.uk/Tools/msa/clustalw2/>.

### 1.3.1 Transcriptional regulation by LTTRs

LTTRs were initially thought to share three main characteristics (Schell, 1993). They respond to exogenous inducer molecules; they bind a consensus operator sequence T-N<sub>11</sub>-A; they are divergently transcribed in relation to the regulator operon (*i.e.* they share a spatial gene organisation in which their encoding gene is located upstream and in the opposite direction of the regulated operon (Figure 1.7)). Divergently transcribed transcriptional units are believed to be more stable during genetic rearrangement and recombination and are believed to have advantages over non-divergent transcription as it only requires the maintenance of one protein to regulate two sets of genes (Beck and Warren, 1988; Reen et al., 2013a).

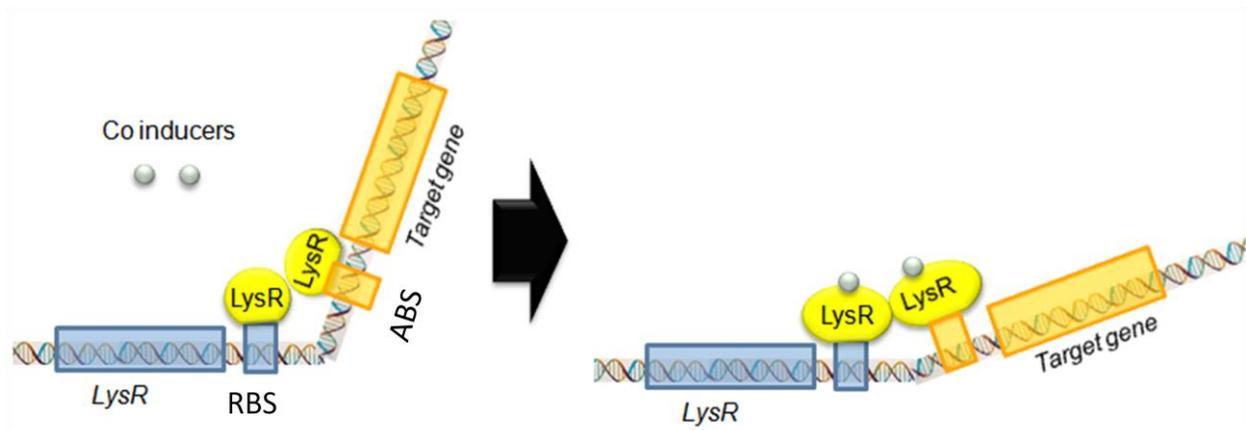
Later studies reviewed by (Maddocks and Oyston, 2008; Momany and Neidle, 2012) also showed that LTTRs usually associate as tetramers (Ezezika et al., 2007; Monferrer et al., 2010; Muraoka et al., 2003; Smirnova et al., 2004), bind to DNA through a wHTH domain (Alanazi et al., 2013a; McFall et al., 1997; Xu et al., 2001) and are activated via the binding of inducer molecules to a C-terminal domain (Ezezika et al., 2007; Monferrer et al., 2010; Smirnova et al., 2004). However, LTTRs which do not strictly follow these characteristics have been identified. Examples here include *Neisseria Meningitidis* OxyR which is activated by the breakage, caused by oxidative stress, of a disulphide cysteine bridge, rather than the binding of an inducer molecule (Choi et al., 2001), the newly identified *Pseudomonas aeruginosa* PA2206, which was shown not to be divergently transcribed (Reen et al., 2013b) and *Neisseria Meningitidis* CrgA which was shown to associate as an octamer (Section 1.3.5) (Sainsbury et al., 2009).

LTTRs bind to two sites on DNA; the Repression (or Regulating) Binding Site (RBS) and the Activating Binding Site (ABS) (Figures 1.7, 1.8). The RBS is a 20-30 bp motif containing a ATAC-N<sub>7</sub>-GTAT repeat with the consensus LTTR binding motif T-N<sub>11</sub>-A located around 65 bases upstream of the *LysR* gene (Schell, 1993). LTTRs bind the RBS in both the presence and absence of an inducer molecule allowing it to serve as an anchoring point (Schell, 1993). The repression of transcription of the *Leers* gene has been seen to be slightly lowered in the presence of an inducer molecule (Coco et al., 1993; van der Meer et al., 1991) but this does not however seem to be a general trend (Tropel and van der Meer, 2004).



**Figure 1.7: The classical model of LTR dependent transcriptional regulation.** The *ltr* gene is transcribed when no LTR protein is bound to the LTR promoter region, ensuring a basal level of LTR (yellow) within the cell. When a LTR is bound to the RBS it inhibits auto expression. The LTR binds upstream of the promoter region of the divergently transcribed target gene and upon the binding of an inducer molecule it activates transcription of the target gene. Figure adapted from (Maddocks and Oyston, 2008).

The ABS contains no consensus sequence but was shown during studies of the binding of ClcR to the *clcA* gene, to shift from position -37 to -41 in the presence of an inducer molecule (McFall et al., 1997). LTR tetramers bind both DNA sites simultaneously and, unlike other transcription factors, regulate transcription by undergoing conformational changes without changing their binding affinity. When binding to DNA in their inactive state LTRs induce a bend in the DNA of the promoter region between  $50^\circ$  and  $100^\circ$ , preventing the binding of RNA polymerase. When a LTR is activated this bend in the DNA relaxes to between  $9^\circ$  and  $50^\circ$  exposing the promoter to the solvent, allowing RNA polymerase recruitment and thus transcription of the regulated gene (Figure 1.8) (Fisher and Long, 1993; Hryniewicz and Kredich, 1995; van Keulen et al., 1998; Parsek et al., 1995; Piñeiro et al., 1997; Toledano et al., 1994; Wang et al., 1992).



**Figure 1.8: A schematic showing the current model of LTTR transcriptional regulation.** Two dimers of a LTTR bind to the RBS and ABS sites of the promoter region and associate to form a tetramer. This results in a bend in the DNA of between  $50^\circ$  and  $100^\circ$ . Upon activation, the LTTR homotetramer undergoes a conformational change which causes a relaxation of the DNA bend, allowing RNA polymerase to bind. Figure adapted from (Maddocks and Oyston, 2008).

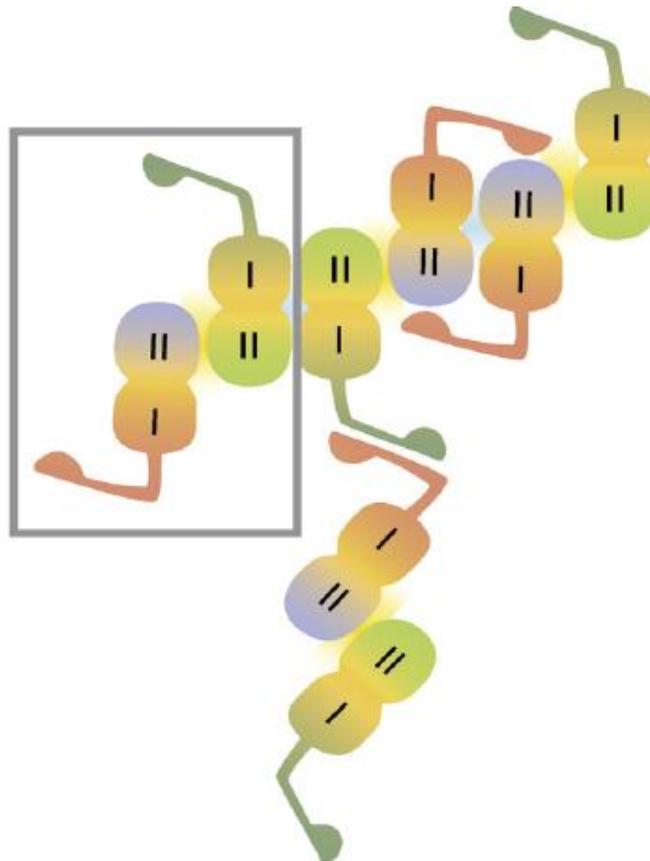
### 1.3.2 The structure of LTTR proteins

LTTRs comprise approximately 330 residues and generally form homotetramers in solution (Ezezika et al., 2007; Monferrer et al., 2010; Muraoka et al., 2003; Smirnova et al., 2004). In LTTR monomers a large C-terminal region contains the Inducer Binding/Switch Domain (IBD) while approximately 60 amino acids in the N-terminal domain forms the winged Helix Turn Helix (wHTH) region required for DNA binding (Momany and Neidle, 2012). The two domains are connected by a flexible linker helix region. The IBC consists of two Rossmann fold like domains (RD1 and RD2) with the inducer binding cleft found (IBC) at their interface. A schematic of the domain structure of LTTR monomers is shown in Figure 1.9.



**Figure 1.9: The domain structure of LTTR monomers.** The N-terminal DNA Binding Domain (DBD) is connected to the C-terminal IBD through a flexible linker helix.

The *in vitro* production of LTR proteins is often made difficult by solubility issues (Ezezika et al., 2007; Xu et al., 2012), most likely caused by the vast array of oligomerisation schemes they can adopt at higher concentrations (Figure 1.10) (Ruangprasert et al., 2010). These can be overcome by truncating the protein leaving only the IBD (Momany and Neidle, 2012). Many structural studies of LTRs have thus focused on the IBDs only. Table 1.2 shows the crystal structures obtained for LTR proteins to date.

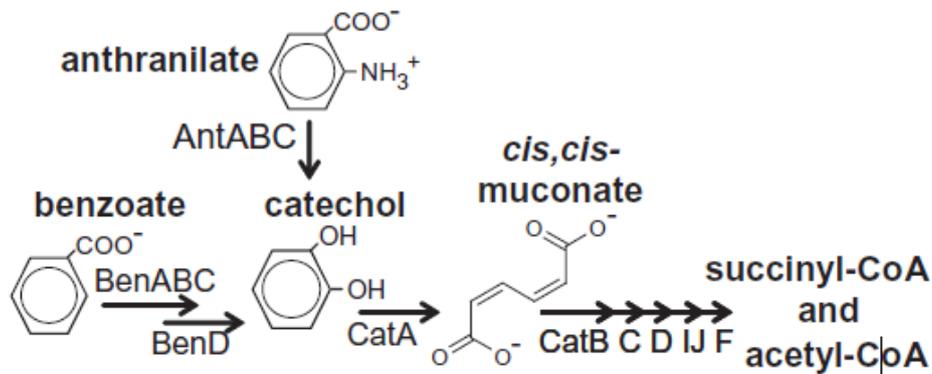


**Figure 1.10: Proposed oligomerisation schemes for the LTR BenM.** RD1 and RD2 of the IBD are denoted I and II. Multiple DntR monomers may have contacts through RD2 and thus create a chain of tetramers. The N-terminal domain of a monomer in one chain may also associate with the N-terminal domain of a different chain. Figure from (Ruangprasert et al., 2010).

Name (PDB ID)	IBD only	Full length protein	Reference
<b>AphB:</b>			
Full length (3SZP)		X	(Taylor et al., 2012)
Full length N100E (3T1B)		X	(Taylor et al., 2012)
<b>BenM:</b>			
Inducer binding domain at high pH (2F8D)	X		(Ezezika et al., 2007)
Inducer binding domain (2F6G, 2F6P)	X		(Ezezika et al., 2007)
Inducer binding domain in complex with benzoate (2F78)	X		(Ezezika et al., 2007)
Inducer binding domain in complex with muconate (2F7A)	X		(Ezezika et al., 2007)
Inducer binding domain R156H, T157S (2H9B)	X		(Craven et al., 2009)
DBD domain in complex with DNA (4IHT)			(Alanazi et al., 2013a)
<b>CatM:</b>		X	(Muraoka et al., 2003)
Inducer binding domain (2F7B)	X		(Ezezika et al., 2007)
Inducer binding domain in complex with muconate (2F7C)	X		(Ezezika et al., 2007)
Inducer binding domain R156H (3FLB)	X		(Craven et al., 2009)
Inducer binding domain V158M (2H98)	X		To Be Published
<b>CbnR:</b> Full length (1IZ1)		X	(Muraoka et al., 2003)
<b>CysB:</b>			
Cofactor binding fragment (1AL3)	X		(Tyrrell et al., 1997)
Inducer binding domain (4GWO)	X		To be Published
<b>CynR:</b> Full length (3HFU)		X	To be Published
<b>DntR:</b>			
Full length In complex with thiocyanate (1UTH)		X	(Smirnova et al., 2004)
Full length In complex with acetate (1UTB)		X	(Smirnova et al., 2004)
Mutants F111L and Y110S,F111V (2UFY, 2UYE)		X	(Lönneborg et al., 2007)
Inducer binding domain (2Y7R)	X		(Devesse et al., 2011)
Inducer binding domain complex with salicylate (2Y7P)	X		(Devesse et al., 2011)
<b>MetR</b> Inducer binding domain (4AB5)	X		(Sainsbury et al., 2012)
<b>OxyR:</b>			
Oxidised form (1I6A)	X		(Choi et al., 2001)
Reduced form (1I69)	X		(Choi et al., 2001)
Reduced form from <i>N. Meningitidis</i> (3JV9)	X		(Sainsbury et al., 2010)
From <i>porphyromonas gingivalis</i> (3HO7)	X		(Svintradze et al., 2013)
<b>PA01:</b> Inducer binding domain (3FZV)		X	To be Published
<b>RovM:</b> Inducer binding domain (3ONM)	X		(Quade et al., 2011)
<b>TsaR:</b>			
In complex with p-toluenesulfonate (3FXU)		X	(Monferrer et al., 2010)
Inducer binding domain (3N6T)	X		(Monferrer et al., 2010)
Inducer binding domain in complex with p-toluenesulfonate (3FXU)	X		(Monferrer et al., 2010)
<b>Un named</b> (3KN3)	X		To be Published

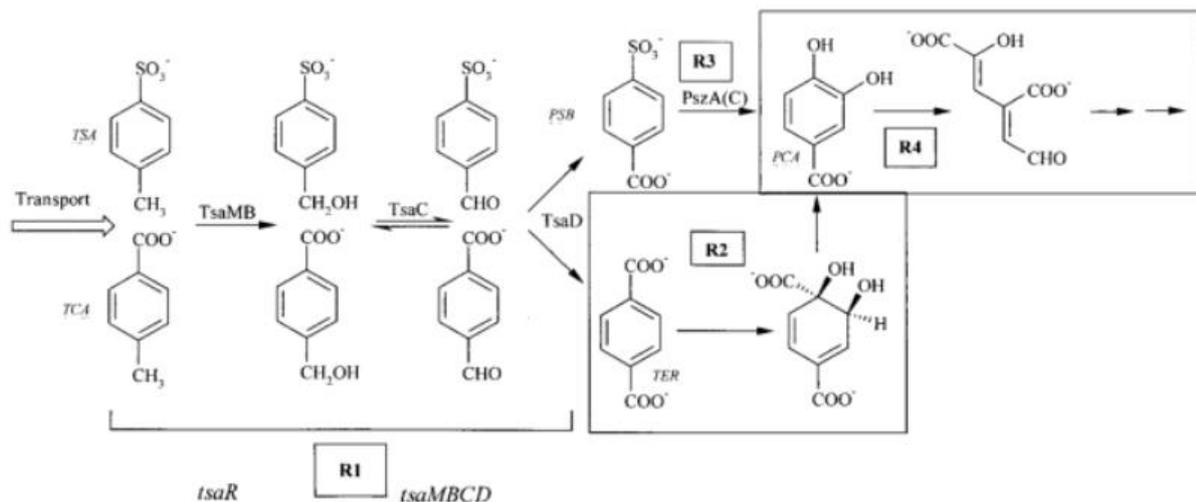
**Table 1.2: Available crystal structures of LTR proteins.** The currently available crystal structures of LTRs. PDB Ids are shown in parenthesis.

As can be seen in Table 1.2, BenM and TsaR are two particularly well studied LTRs. BenM is a transcription factor from the soil bacterium *Acinetobacter baylyi* in which it regulates the transcription of genes, *BenABCD*, involved in the degradation of benzoate (Figure 1.11) (Collier et al., 1998).



**Figure 1.11: The pathway for the degradation of benzoate and anthranilate in the soil bacterium *Acinetobacter baylyi*.** BenM provides transcriptional regulation for all genes responsible for the breakdown. Figure as published by (Craven et al., 2009).

TsaR regulates the *tsa* operon containing the *tsaMBCD* pathway responsible for the degradation of *p*-toluenesulfonate (TSA) and *p*-toluenecarboxylate (TCA) in *Comamonas Testosterone* (Figure 1.12) (Tralau et al., 2003).



**Figure 1.12: The pathway for the degradation of TSA and TCA in *Comamonas Testosterone*.** TsaR provides transcriptional regulation for genes within the box R1 that are responsible for the breakdown of TSA and TCA. Figure as published by (Tralau et al., 2003).

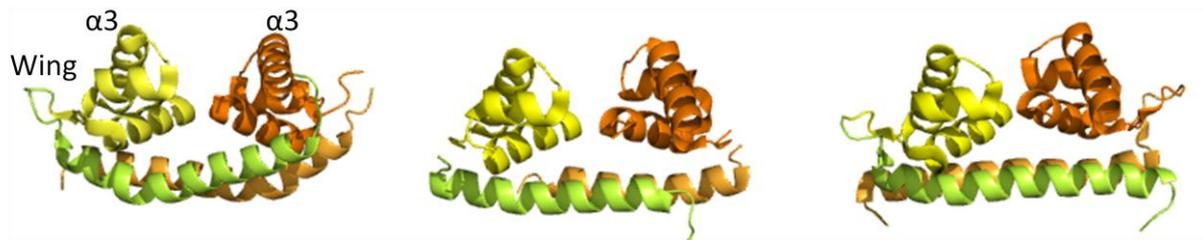
### 1.3.3 The LTTR DNA Binding Domain.

In the crystal structures of full-length LTTRs the DBD/linker helix domains of the LTTR monomers associate to form wHTH dimers that flank a central IBD tetrameric core (Figure 1.13). The wHTH motif consists of three helices noted  $\alpha 1$ ,  $\alpha 2$  and  $\alpha 3$ . Upon contact with DNA the outer helix,  $\alpha 3$ , also known as the recognition helix, is embedded into the major groove of its target site and allows for the two remaining helices to have contact with the more distal parts of the DNA (Brennan, 1993). As shown in Table 1.3 and Figure 1.14 the DBD/linker helix domains of different LTTRs display a higher degree of sequence similarity than is observed for the full length proteins (see Table 1.1).

Recently two crystal structures of the BenM DBD/linker helix domain bound to DNA were solved (Alanazi et al., 2013a) (Figure 1.15). As can be seen from Figure 1.15, in these crystal structures the DBD/linker helix dimers are very similar in structure to those shown in Figure 1.14, and bind DNA in such a way that the two recognition helices in the wHTH dimer occupy two consecutive turns of the major groove of the DNA. This is unlike what is observed for other transcriptional regulators where the two recognition helices of a dimer occupy a continuous region of the major groove (Xu et al., 2001). The crystal structures solved by (Alanazi et al., 2013a) described complexes with two different DNA fragments; the promoter regions for the two divergently transcribed genes *benA* (T-N<sub>11</sub>-A site: ATACTCCATAGGTAT) and *catB* (T-N<sub>11</sub>-A site: ATACCTTTTATAGTAT) and show only minor differences. In both cases 27 amino acids of the BenM dimer make 36 contacts with the DNA duplex. The major interactions seen are between recognition helix  $\alpha 3$  and the major grooves of the ATAC sites at either end of the DNA duplex while the wings of the wHTH domain interact with the minor groove.

The crystal structure also shows that the recognition of the ATAC motifs is highly dependent on the conserved residue Ser33 (see Figure 1.15, bottom) which interacts with the methyl group of the thymidine base through its C <sup>$\beta$</sup>  atom and forms, via its hydroxyl group, hydrogen bonds with the phosphate group of the ATAC 5' adenosine residue. As can be seen in Figure 1.14 S33 is generally conserved throughout LTTRs with its hydrogen binding properties maintained in CbnR and AphB where it is replaced with Threonine. The importance of S33 for recognition of the ATAC site is underlined by the fact that the S33N mutation in OxyR

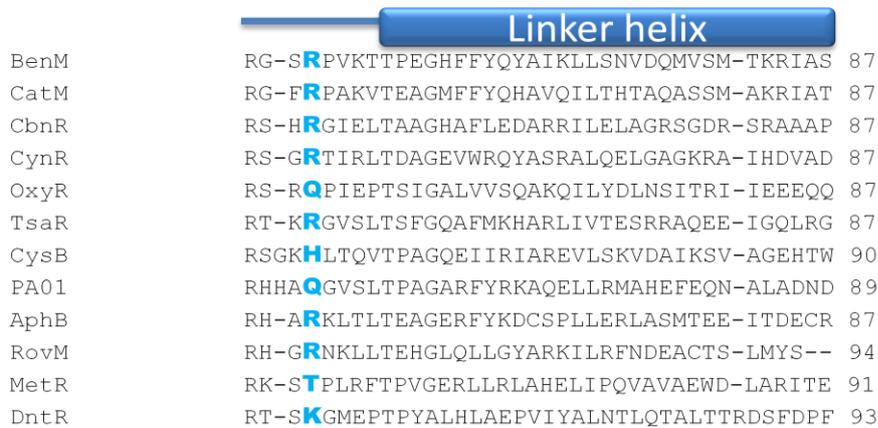
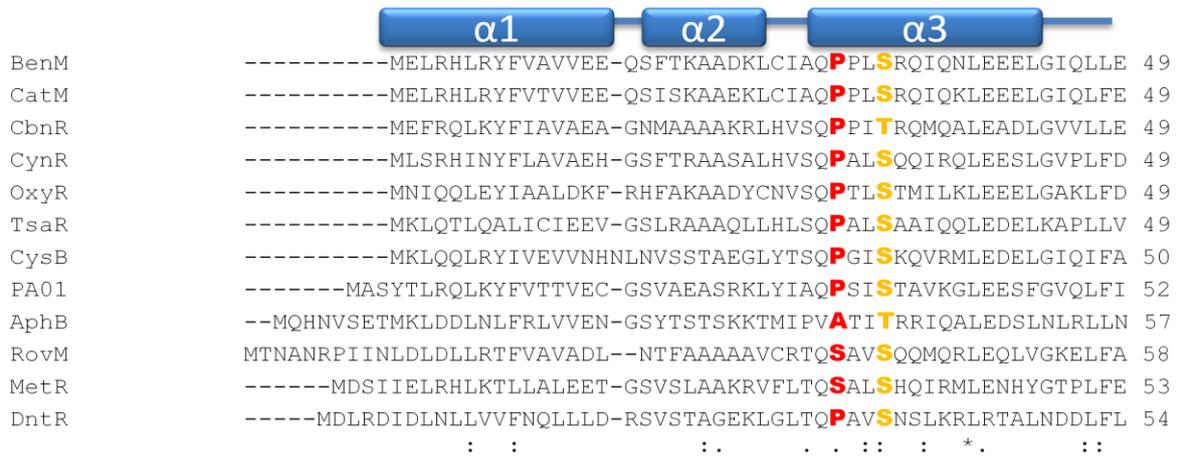
(Kullik et al., 1995) and corresponding mutations S34R in CysB (Colyer and Kredich, 1994) and S38P in GcvA (Jourdan and Stauffer, 1998) have been shown to abolish the DNA binding properties of these LTTRs. Other important residues in DNA recognition are P30, R34 and R53 (Alanazi et al., 2013a) which are all also well conserved amongst LTTRs (Figure 1.14). Based on the conservation of the critical amino acids in the BenM DBD it seems likely that the model of DNA binding observed for BenM is shared amongst all LTTRs.



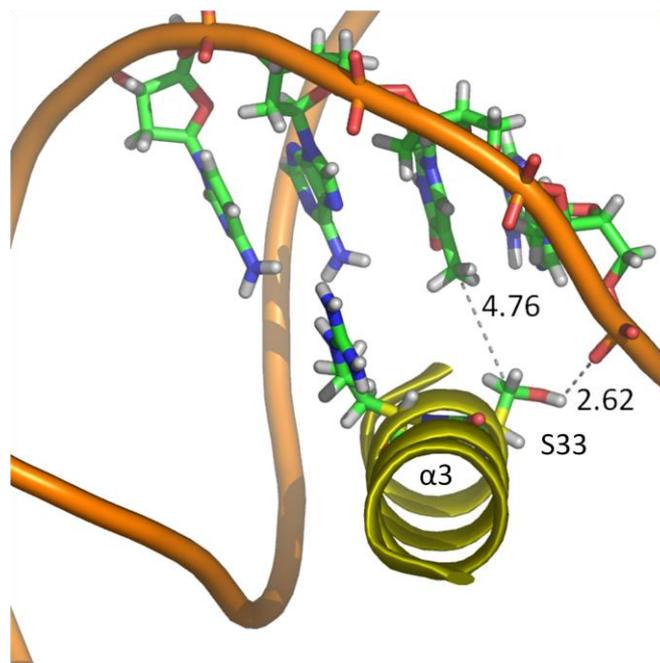
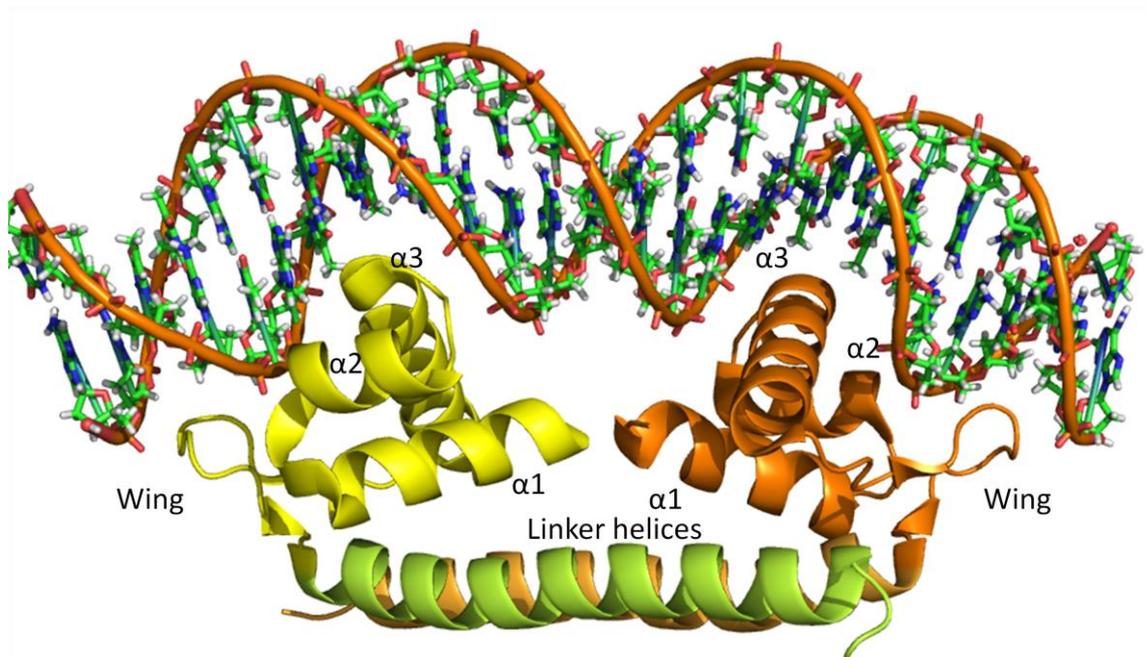
**Figure 1.13: The crystal structure of dimeric DBD/linker helix domains of LTTR homotetramers.** From crystal structures of (from left to right): Tsar (Monferrer et al., 2010), DntR (Smirnova et al., 2004), CbnR (Muraoka et al., 2003).

	BenM	CatM	CysB	PA01	TsaR	AphB	CbnR	CynR	OxyR	RovM	MetR
BenM											
CatM	72.41										
CysB	37.21	39.53									
PA01	37.93	39.08	32.18								
TsaR	31.03	33.33	26.74	26.44							
AphB	32.18	31.03	22.09	30.00	28.74						
CbnR	39.08	37.93	27.91	34.48	35.63	31.03					
CynR	40.23	37.93	29.07	31.03	33.33	29.89	42.53				
OxyR	31.03	32.18	27.91	25.29	27.59	19.54	29.89	35.63			
RovM	32.14	32.14	26.51	24.14	25.00	18.48	30.95	30.95	20.24		
MetR	35.63	34.48	26.74	28.89	27.59	19.78	26.44	32.18	21.84	29.55	
DntR	21.84	22.99	18.39	22.22	20.69	14.13	18.39	19.54	19.54	24.72	24.18

**Table 1.3: Sequence identity matrix of the N-terminal of LTTR proteins.** The alignment was performed using the EMBL Clustalw2 server <http://www.ebi.ac.uk/Tools/msa/clustalw2/>.



**Figure 1.14: Amino acid sequence alignment of the DBD/linker helix domains of various LTRs.** The regions forming the three  $\alpha$ -helices in the wHTH bundle and the linker helix as determined by (Alanazi et al., 2013a) are shown. The alignment was performed using the EMBL Clustalw2 server <http://www.ebi.ac.uk/Tools/msa/clustalw2/>. Residues corresponding to BenM P30 (red), S33 (yellow) and R53 (blue) are indicated.



**Figure 1.15: The binding of a dimer of the BenM DBD linker helix domain to its promoter region DNA.** Top: The crystal structure of the BenM dimeric DBD/linker helix domain bound to DNA as outlined in the main text. The recognition helices of the wHTH dimer interact with consecutive turns of the major groove of the DNA. Bottom: Specific interactions between  $\alpha 3$  and the ATAC site as seen in the crystal structure reported by (Alanazi et al., 2013a).

### 1.3.4 LTTR Inducer Binding Domains

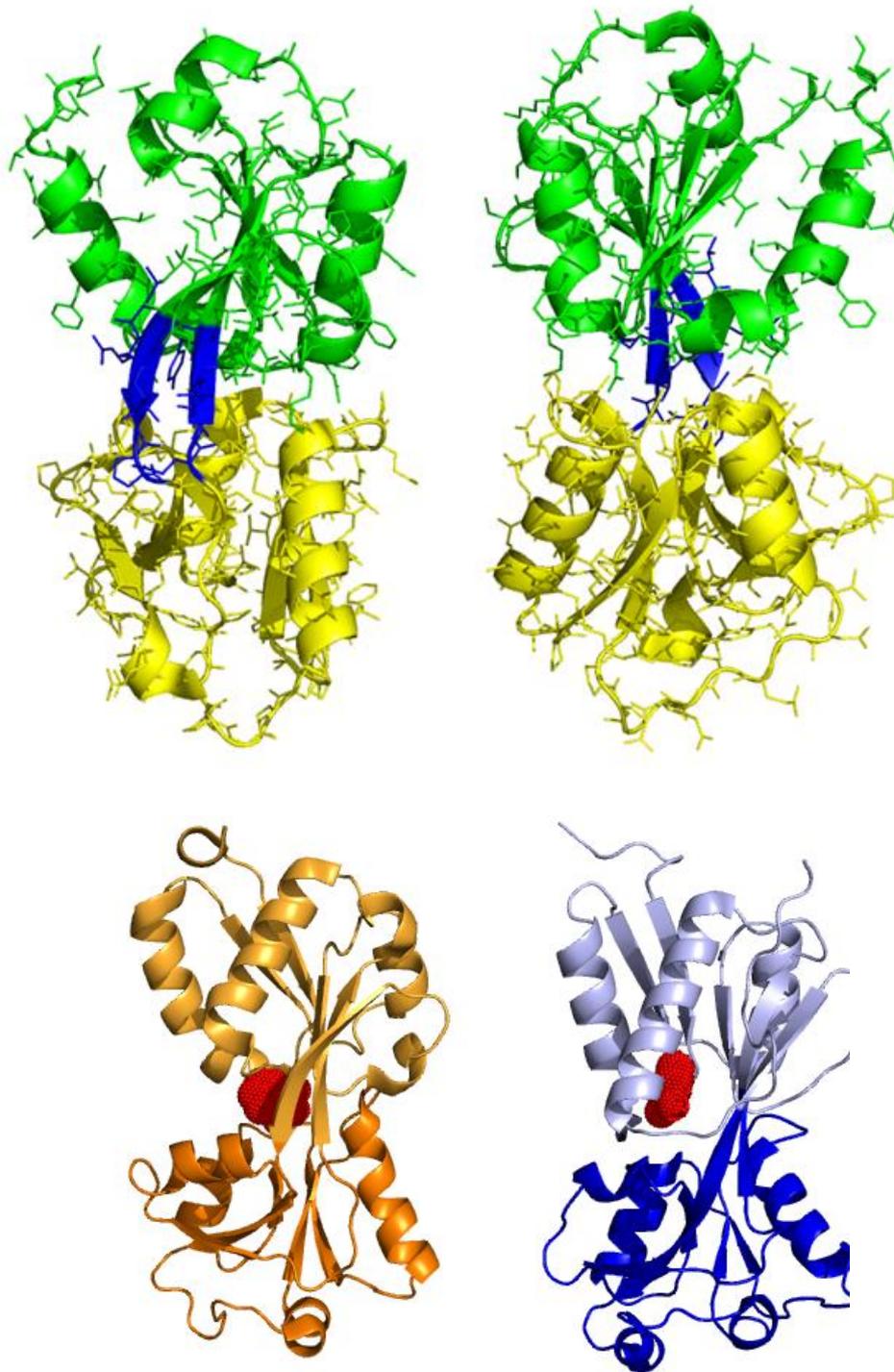
The three dimensional structures of LTTR Inducer Binding Domains (IBDs) are well characterized and in spite of a relative lack of sequence identity (Table 1.4, Figure 1.16) they are structurally very similar (Craven et al., 2009; Devesse et al., 2011; Ezezika et al., 2007; Lönneborg et al., 2007; Muraoka et al., 2003; Sainsbury et al., 2012; Smirnova et al., 2004; Taylor et al., 2012; Tyrrell et al., 1997) comprising two globular Rossmann-Fold like subdomains, denoted RD1 and RD2 linked by a hinge region (Figure 1.17, bottom).

For LTTRs activated by inducer molecules, a conserved inducer binding site (IBC) is located between the two RD subdomains of the IBD close to the hinge region. The crystal structures of inducer-bound IBDs of DntR (Devesse et al., 2011) and BenM (Craven et al., 2009) are shown in Figure 1.17, bottom.

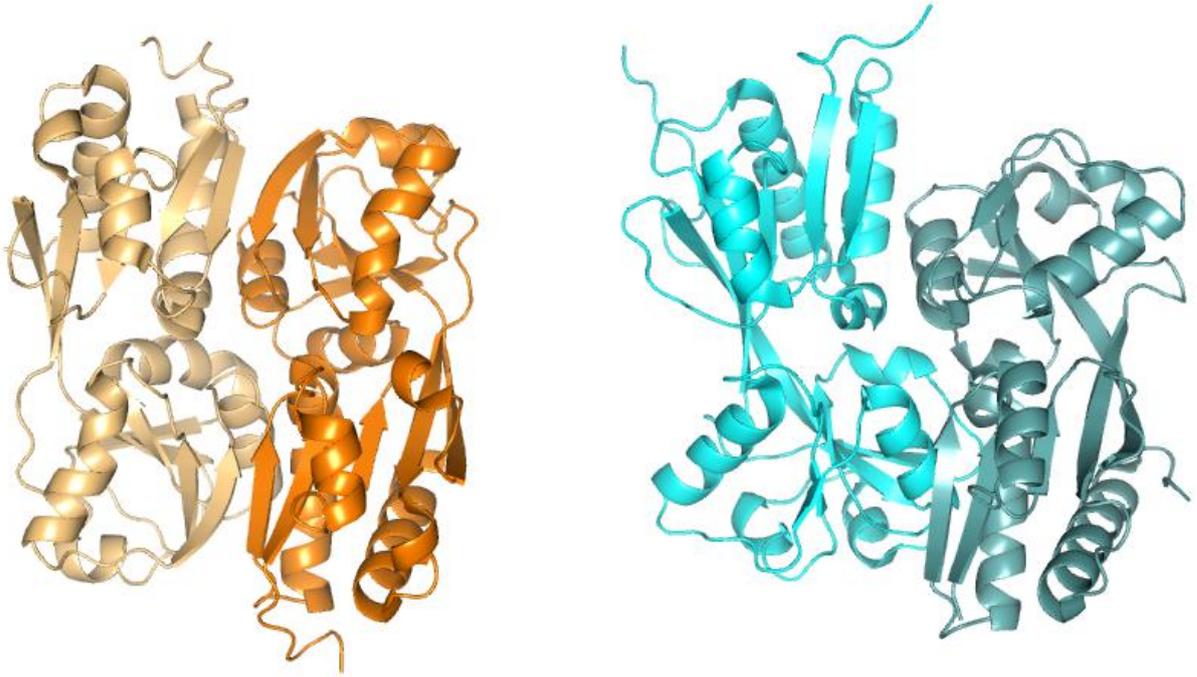
	BenM	CatM	CbnR	PA01	AphB	CynR	TsaR	CysB	OxyR	RovM	MetR
BenM											
CatM	53.24										
CbnR	25.24	23.79									
PA01	21.33	20.38	17.16								
AphB	16.42	14.93	12.06	13.86							
CynR	16.59	18.48	17.16	18.57	13.2						
TsaR	17.56	14.63	15.27	15.94	12.76	20.98					
CysB	11.27	12.74	14.15	19.16	10.95	22.64	17.79				
OxyR	15.49	11.79	15.61	18.69	10.5	16.51	15.31	16.2			
RovM	14.29	15.84	13.33	17.73	14.72	20.5	17.01	17.21	18.81		
MetR	15.17	17.14	13.3	19.53	13	15.71	13.11	17.89	15.35	14.01	
DntR	11.71	10.24	13.37	11.96	10.24	12.32	14.85	12.56	14.56	13.13	13.11

**Table 1.4: Sequence identity matrix of the C-terminal of LTTR proteins.** The alignment was performed using the EMBL Clustalw2 server <http://www.ebi.ac.uk/Tools/msa/clustalw2/>.





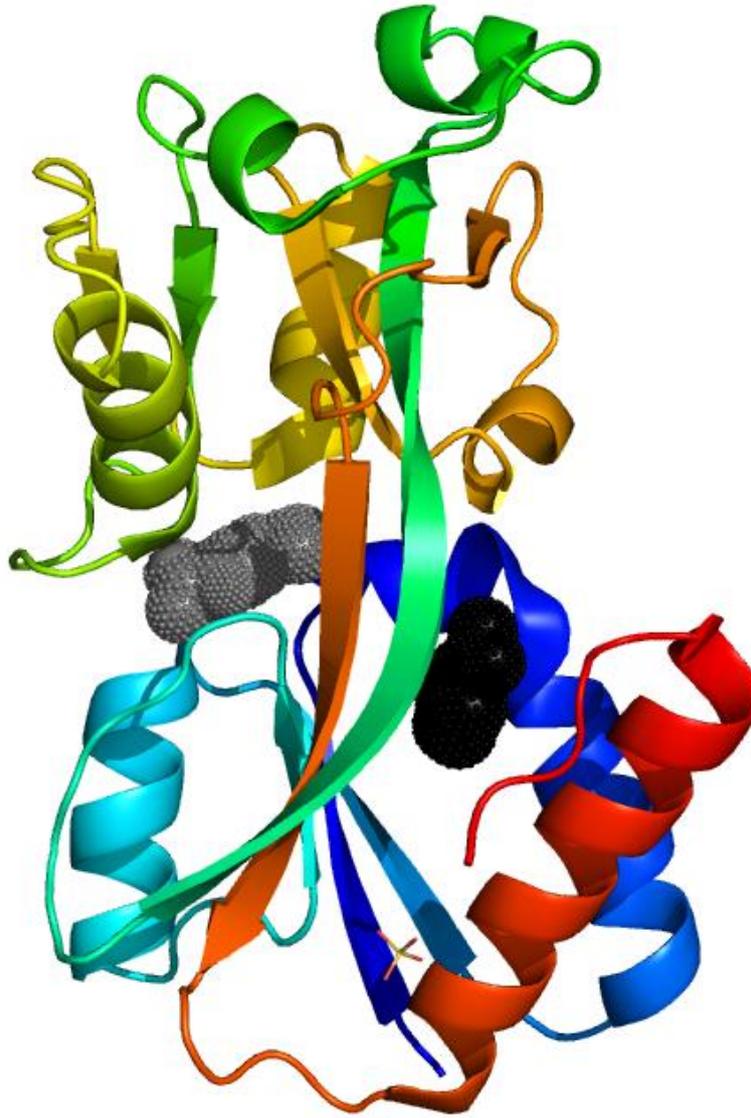
**Figure 1.17: The crystal structures of the LTTR IBD monomer.** Top: Orthogonal views of the structure of a LTTR IBD monomer. RD1 is shown in green, RD2 shown in yellow and the linker region shown in blue. The IBD shown is that seen in the crystal structure of RovM from (Quade et al., 2011). Bottom: Inducer-bound IBDs as seen in the crystal structures of the IBDs of DntR (left) (2Y7K) and BenM (right) (2F7A). For both proteins residues 90-300 are shown. The inducer molecules salicylate and muconate respectively are shown in red and are bound in the conserved IBC. The two Rossmann-like fold domains for each IBD are shown in different nuances.



**Figure 1.18: The crystal structures of LTR IBD dimers.** From left to right: DntR (2Y7K) and BenM (2F7A). Residues 90-300 are shown. The two monomers comprising each dimer are shown in different nuances.

In crystal structures of constructs consisting only of LTR IBDs, two IBD monomers associate head-to-tail to form dimers (Devesse et al., 2011; Ezezika et al., 2007; Quade et al., 2011; Svintradze et al., 2013; Tyrrell et al., 1997) as depicted in Figure 1.18.

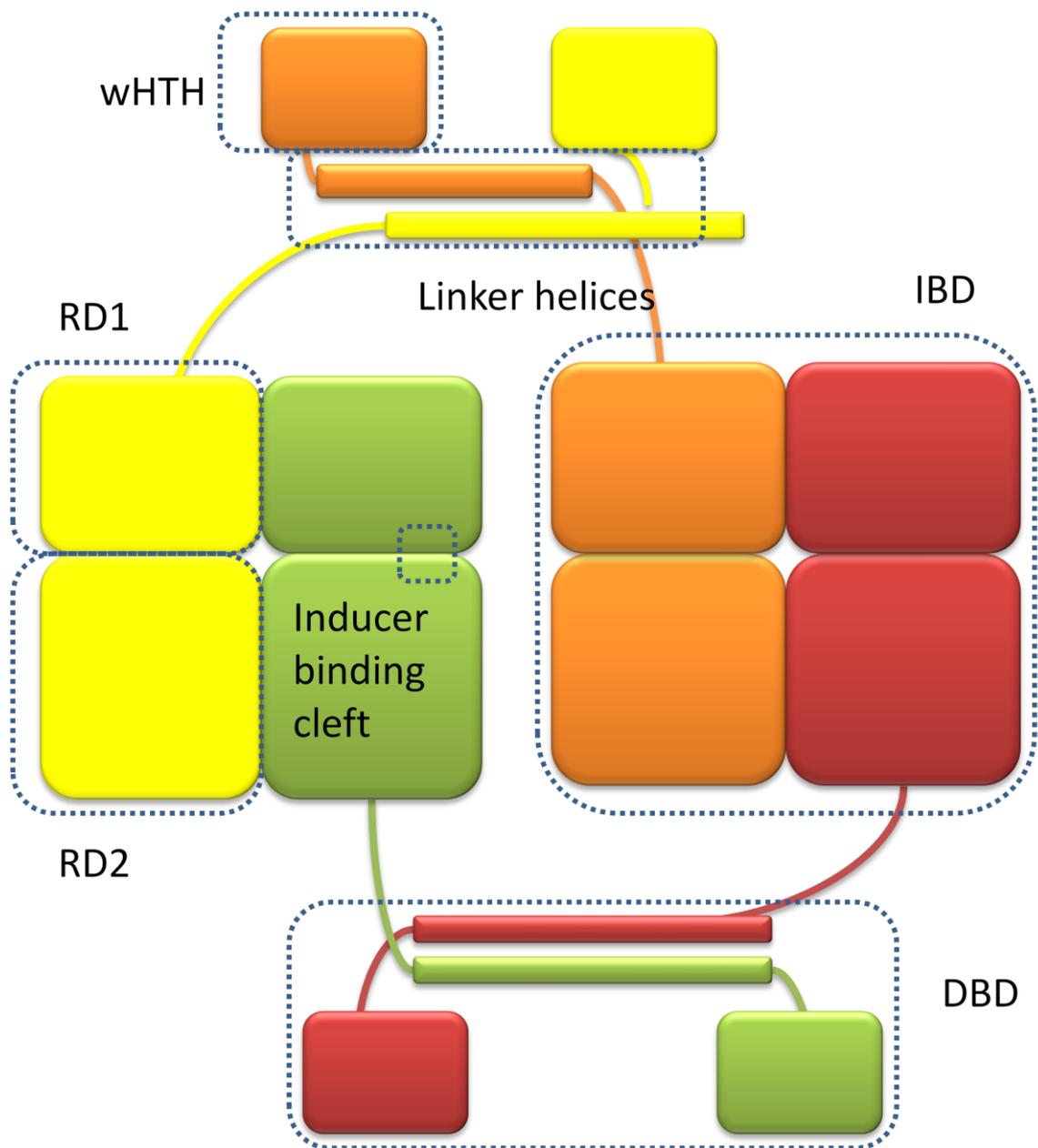
However, in spite of the strong structural similarities shown in Figure 1.17 and 1.18, there are differences in how inducer molecules are bound by LTR IBDs. Indeed, some show effector binding in sites other than the IBC. BenM activates gene expression both in the presence of benzoate and muconate and when both molecules are present they have a synergistic effect on transcriptional regulation, higher than their singular additive effects (Bundy et al., 2002). Structural studies of BenM identified muconate in the conserved LTR IBC but showed benzoate to bind in a previously unobserved binding site as shown in Figure 1.19 (Ezezika et al., 2007). When benzoate is bound in this site it enhances, through electrostatic interactions, the effect of muconate binding, thus explaining the synergistic effect of the two molecules on transcriptional regulation.



**Figure 1.19: The crystal structure of a BenM monomer with both benzoate and muconate-bound.** Benzoate is shown in black spheres and muconate is shown in grey spheres. As published by (Ezezika et al., 2007).

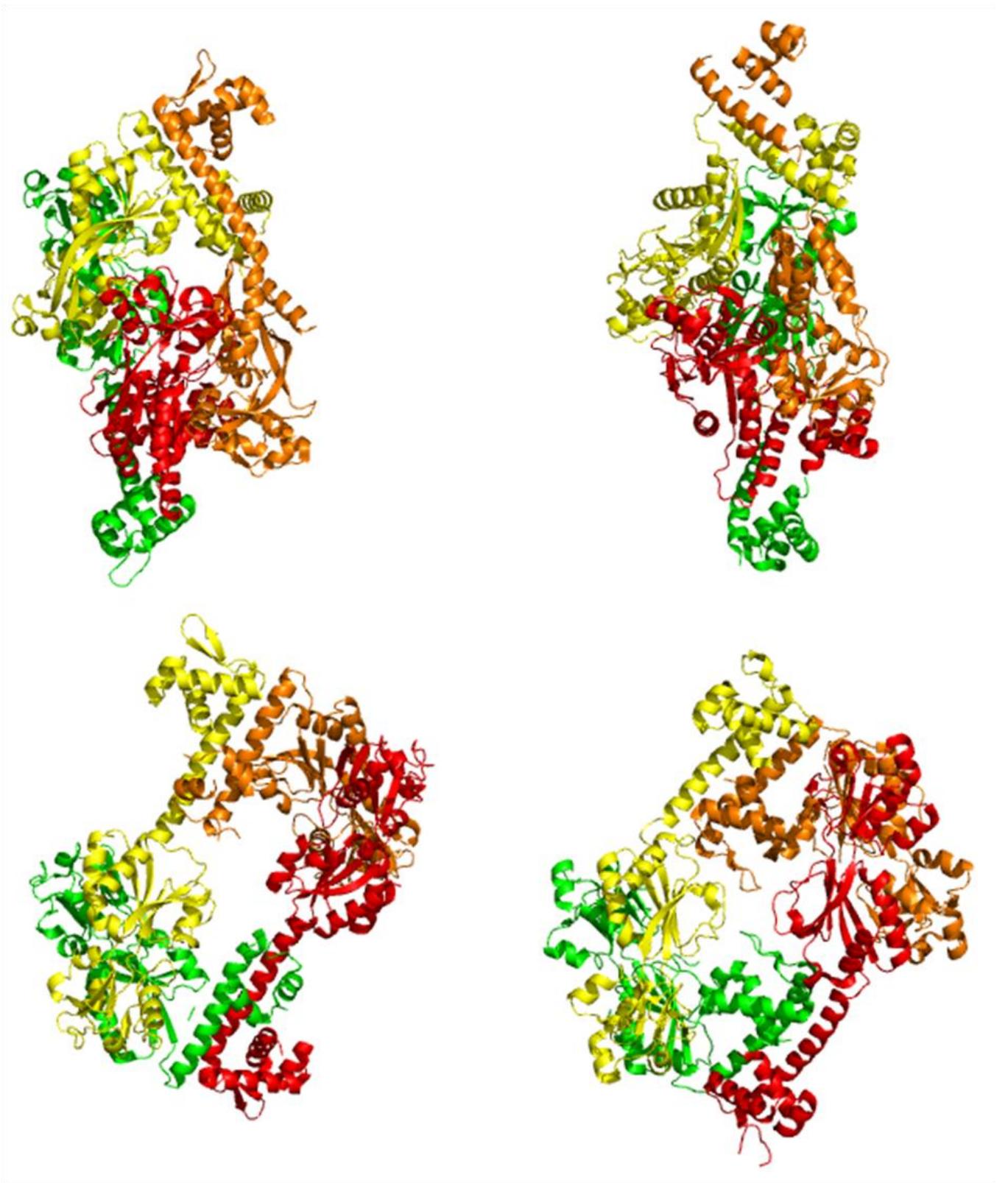
### 1.3.5 The crystal structure of full length LTTRs

Despite difficulties in expression, purification and crystallisation caused by a tendency to aggregate (Section 1.3.2, Figure 1.10) (Ezezika et al., 2007; Xu et al., 2012), the crystal structures of some full length LTTRs have been obtained (Table 1.2). These generally show LTTRs to associate as homotetramers with a quaternary structure comprising a tetrameric core (resulting from the association of two “head-to-tail” IBD dimers) that are connected “head-to-head” through via the dimerisation of two DBD/linker helix domains (Figure 1.20).



**Figure 1.20: Schematic of the crystal structure of a typical LTR homotetramer.** A LTR tetrameric dimer of dimers. Each monomer is shown in different colours. The position of a DNA and Inducer binding domain are indicated.

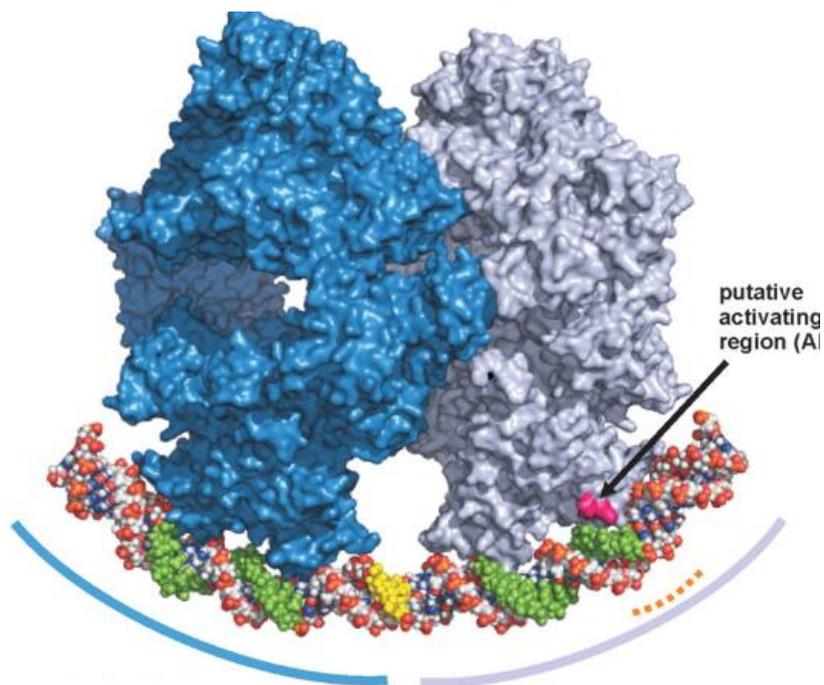
In various crystal structure analyses LTR homotetramers have been observed in either “compact” (Muraoka et al., 2003; Smirnova et al., 2004) or “open” (Monferrer et al., 2010; Taylor et al., 2012) conformations as shown in Figure 1.21. The compact state is maintained through interactions between the C-terminal domains of the head-to-tail dimers. These interactions are abolished in the open state in which the tetrameric state is maintained only through “head-to-head” interactions between the DBD/linker helix domains.



**Figure 1.21: Crystal structures of open and compact full length LTR homotetramers.** Top: CbnR, closed conformation (PDB 1IXC; (Muraoka et al., 2003)) PA01, closed conformation (PDB 3FZV). Bottom: Tsar, open conformation (PDB 3FZJ; (Monferrer et al., 2010)), AphB, open conformation (PDB 3T1B; (Taylor et al., 2012)) Monomers are shown in colours consistent with the colour scheme used in Figure 1.20.

### 1.3.6 Non tetrameric LTTRs

In 2009 the crystal structure of the LTTR CrgA was solved (Sainsbury et al., 2009). In the crystal CrgA presented itself as an octamer, also confirmed to be its oligomeric state in solution by analytical ultracentrifugation. This conformation gives rise to four dimeric DNA-binding wHTH domains. However based on DNA footprinting studies the authors speculate that two octamers associate to form a dimer of octamers, binding DNA with one dimeric wHTH dimer per octamer and leaving the other three unoccupied (Figure 1.22) (Sainsbury et al., 2009). Since the publication of the article describing CrgA no other LTTR has been found to associate as an octamer (Momany and Neidle, 2012). Nevertheless, while CrgA is currently a singular case, its crystal structure shows that some variation may exist within the LTTR protein family.



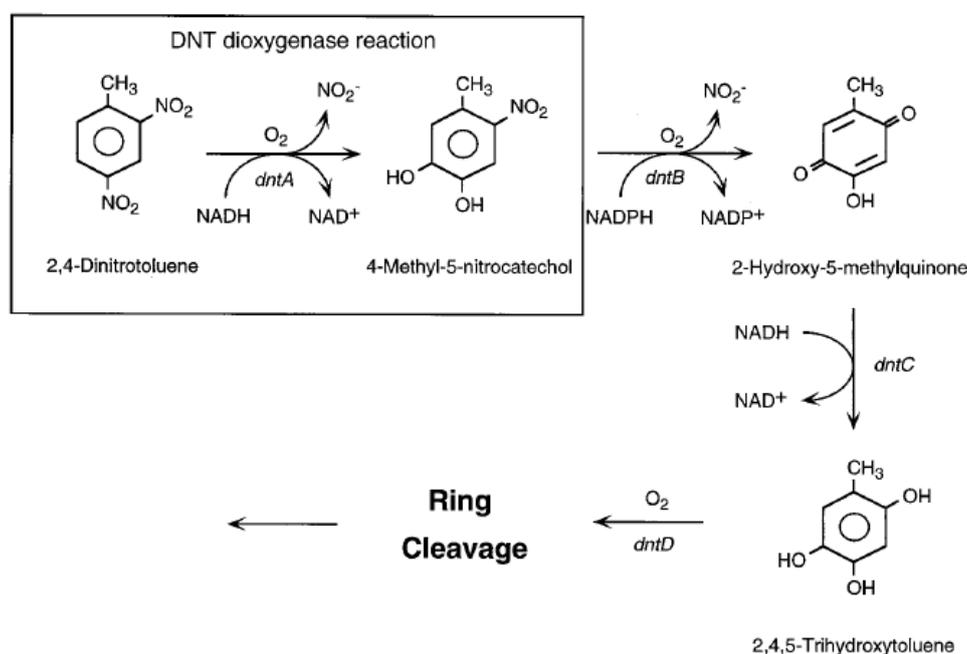
**Figure 1.22:** The proposed oligomerisation and DNA binding scheme of CrgA. Figure as published by (Sainsbury et al., 2009).

### 1.4 The LysR transcriptional regulator DntR

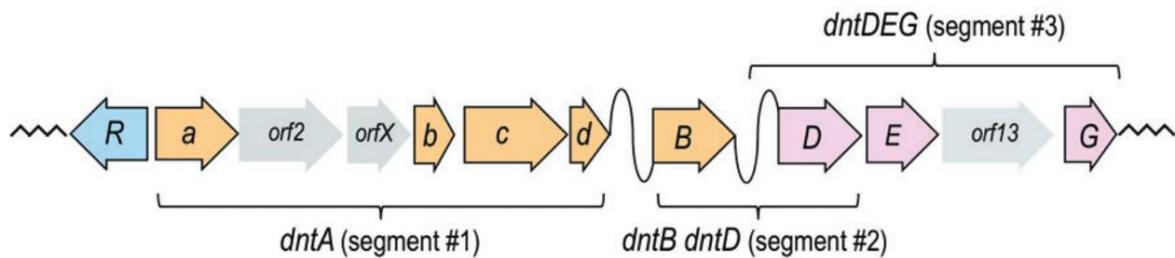
Bacteria that inhabit sites polluted with chemicals that were not introduced into Nature until the onset of synthetic chemistry slowly evolve catabolic pathways for these xenobiotic compounds (Cases and de Lorenzo, 2005). 2,4 dinitrotoluene (DNT) is a synthetic soil

pollutant used in the production of polyurethane foams and in the production of the explosive TNT (Ju et al., 2007). DNTs are highly toxic, converting haemoglobin into methemoglobin, and have a threshold limit value of 1.5 mg/m<sup>3</sup> (Center for Disease Control: <http://www.cdc.gov/niosh/npg/npgd0235.html>).

Since the introduction of DNTs into the environment several species of soil bacteria have evolved pathways which allow them to use DNTs as a carbon and energy source (Johnson et al., 2002; Brüning et al., 2002). The evolution of such a novel degradation pathway requires a bacterium to overcome three problems: the shift of some of its current metabolic enzymes into a new pathway; the evolution of a regulatory device for the expression of the corresponding catalytic genes; the avoidance of stress caused by the substrate itself (Kivisaar, 2009). A bacterium on the evolutionary road to overcoming these problems is *Burkholderia* sp. strain (de las Heras et al., 2011; Suen and Spain, 1993) the genome of which contains genes coding for enzyme pathways that catabolise DNT all the way to pyruvate (Figure 1.23). The most critical step in the degradation process is the deoxygenation of DNT by the enzyme DntA in positions 4 and 5 to yield 4-Methyl-5-nitrocatechol.



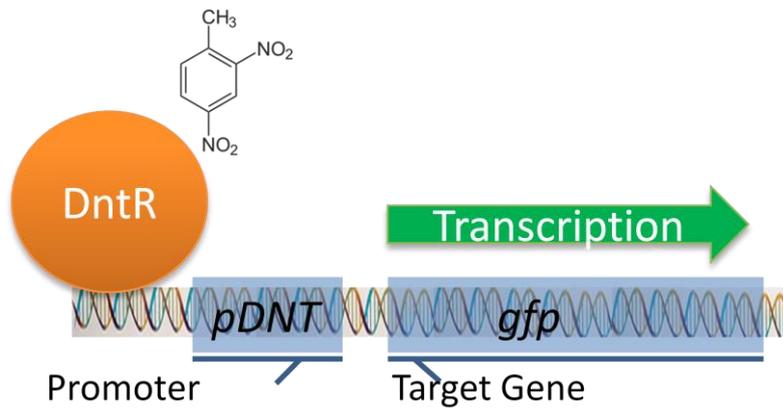
**Figure 1.23: DNT degradation in *Burkholderia* sp.** Clockwise from top left: The DNT deoxygenase (DntA) hydroxylates the aromatic ring of DNT at positions 4 and 5 yielding 4-methyl-5nitrocatechol. A mono oxygenation (DntB) subsequently removes the remaining nitro group and the product reduced to 2,4,5 trihydroxytoluene (DntC). The ring is then cleaved through a CoA-dependent methylmalonate semialdehyde dehydrogenase (DntD). The product eventually converted into pyruvate. Figure from (de las Heras et al., 2011).



**Figure 1.24: The arrangement of the catabolic genes in the DNT degrading pathway in *Burkholderia sp.*** The gene encoding the transcription factor DntR is labelled R and is divergently transcribed of the *dntA* cluster. Figure from (de las Heras et al., 2011).

Having solved the issue of establishing a new catabolic pathway the bacterium also requires a regulation mechanism for the pathway. The organization of genes related to the DNT catabolic pathway in *Burkholderia sp.* is shown in Figure 1.24. The transcription of the most critical component of the pathway is regulated by the LTR DntR, placed divergently of the *dntA* gene. This placement would indicate that DntR will bind DNT as its inducer molecule and subsequently activate the transcription of the *dntA* genes. However, DntR is not activated by DNT but is strongly activated by the small molecule salicylate, an intermediate in the degradation pathway of naphthalene (Lönneborg et al., 2007; Smirnova et al., 2004). This indicates that although *Burkholderia sp.* has evolved the catabolic machinery for the degradation of DNT, the regulation of the pathway still has to be optimised. DntR thus represents an evolutionary snapshot of a TF on the way to proper transcriptional regulation (de las Heras et al., 2011).

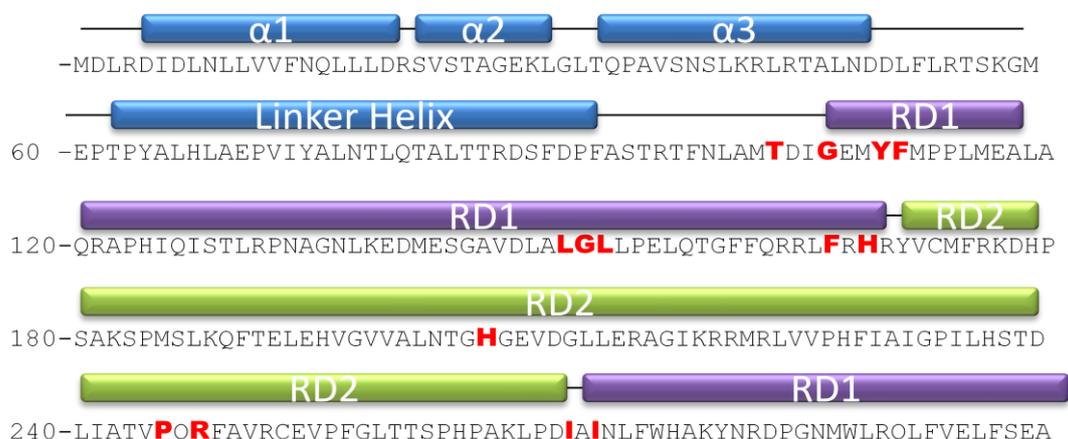
The arrangement of genes as shown in Figure 1.24 prompted studies aimed at using DntR as a biosensor for DNT. The premise of such a biosensor can be seen in Figure 1.25 (Ng and Forsman, 2000). This however will require the production of an engineered DntR with a modified specificity (*i.e.* activated specifically by DNT). While progress has been made in this area (Lönneborg et al., 2007, 2012) the idea of a DntR-based biosensor for DntR is very much at a preliminary stage and thus will not be further discussed here.



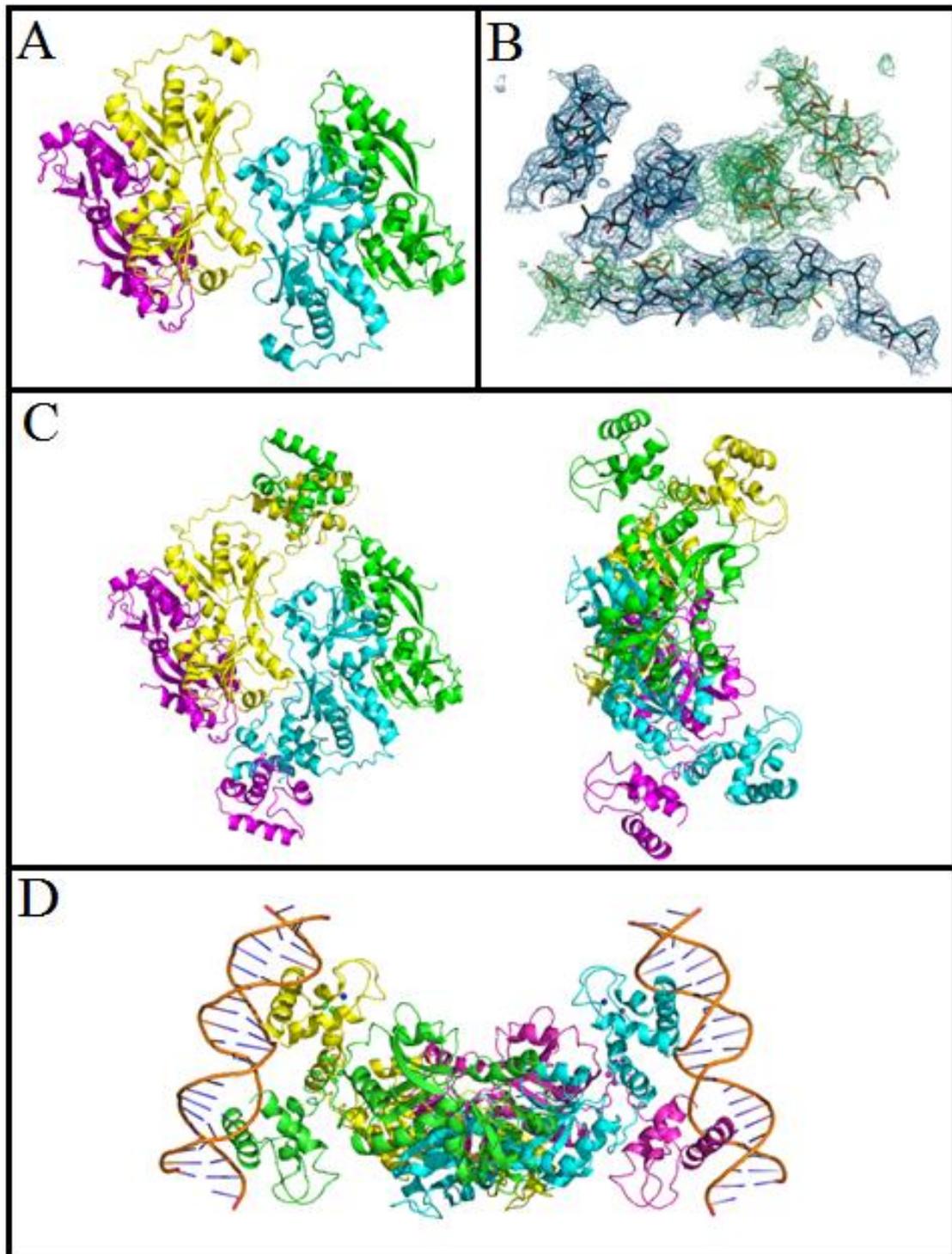
**Figure 1.25: The premise of a biosensor for DNT.** The gene transcription of a fluorescent protein such as *gfp* is placed downstream of the *pDnT* promoter and its transcription is thus under the control of DntR. In the presence of DNT DntR would be activated and allow the transcription of the fluorescent protein. Adapted from (Ng and Forsman, 2000).

#### 1.4.1 The structure of DntR

DntR consists of 301 amino acids in which the 64 residue N-terminal DNA binding wHTH domain is linked to the C-terminal IBD by a linker helix about 20 amino acids in length (Figure 1.26). Mobility shift assays confirmed that DntR binds DNA both in the absence and presence of an inducer molecule, salicylate, and that the binding of this inducer did not significantly change DNA binding affinity (Smirnova et al., 2004). This is compatible with the current model of activation in which LTRs bind DNA both in the presence and absence of an inducer molecule, (see Figure 1.8).



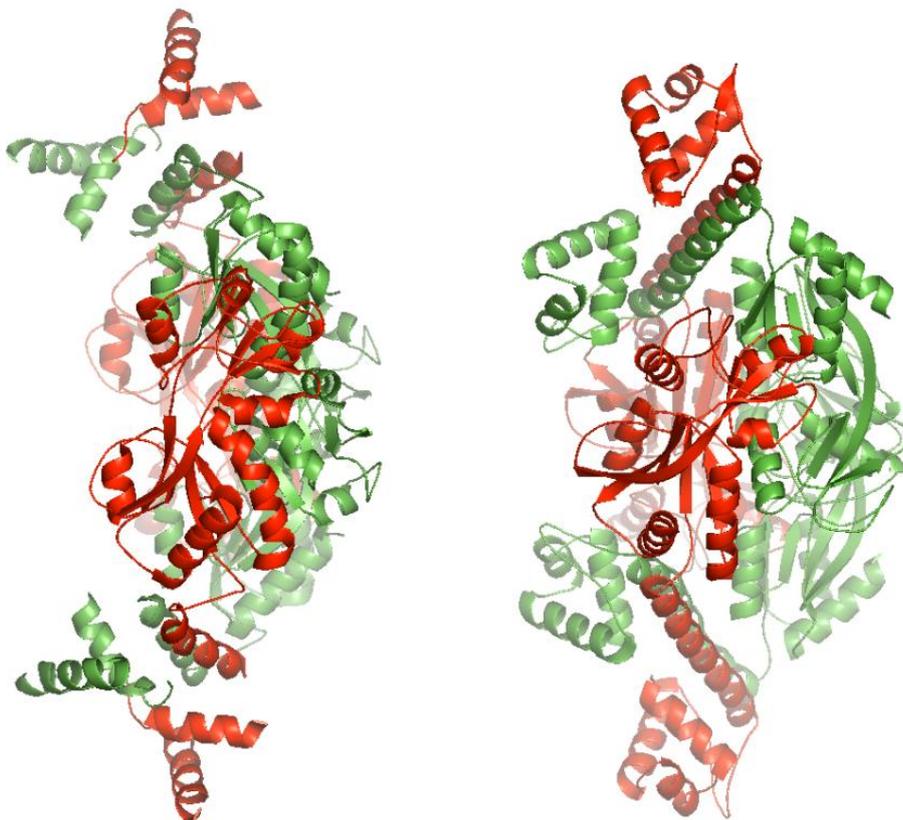
**Figure 1.26: The domain structure of DntR.** As for other LTRs, DntR consists of a wHTH helix motif comprising three  $\alpha$ -helices:  $\alpha 1$ ,  $\alpha 2$  and  $\alpha 3$  joined to the IBD by a linker helix. The residues which form the IBC are shown in red. Domain structure as determined in (Smirnova et al., 2004).



**Figure 1.27: The current available crystal structure of the full length DntR homotetramer.** The monomers are shown in different colors. (A): Structure of the DntR tetrameric core based on the combination of two asymmetric units (PDB 1UTB). The DBD/linker helix domains for all four monomers were disordered. (B): Low level electron density of the DBD linker domains allowed the modelling of the positions of these domains to provide a model for the full length structure of DntR. (C): Two orthogonal views of the model of the structure of homotetrameric structure of full length DntR. (D): A theoretical model of DNA binding to a full length DntR tetramer. Figures from (Smirnova et al., 2004).

The crystal structures of full length DntR in complex with either acetate or thiocyanate were solved in 2004 (Figure 1.27A) (Smirnova et al., 2004). The N-terminal wHTH regions in both crystal structures reported very poor electron density and final models of full-length DntR monomers were based on fitting a homology model into this region for the two DntR molecules found in the asymmetric unit (Figure 1.27B) (Smirnova et al., 2004).

As depicted in Figure 1.27C the crystal structures obtained showed a DntR homotetramer in a compact conformation. Comparison of the crystal structures of DntR and CbnR homotetramers revealed differences in the positions of the N-terminal wHTH domains (Figure 1.28) which were consistent with a relaxation of bound DNA. Based on this, and due to the fact that the IBCs in the crystal structures of DntR were occupied by either thiocyanate or by acetate, it was speculated that both acetate and thiocyanate mimicked the binding of salicylate in the IBC and that the crystal structure of DntR was that of a LTR protein in its active state.

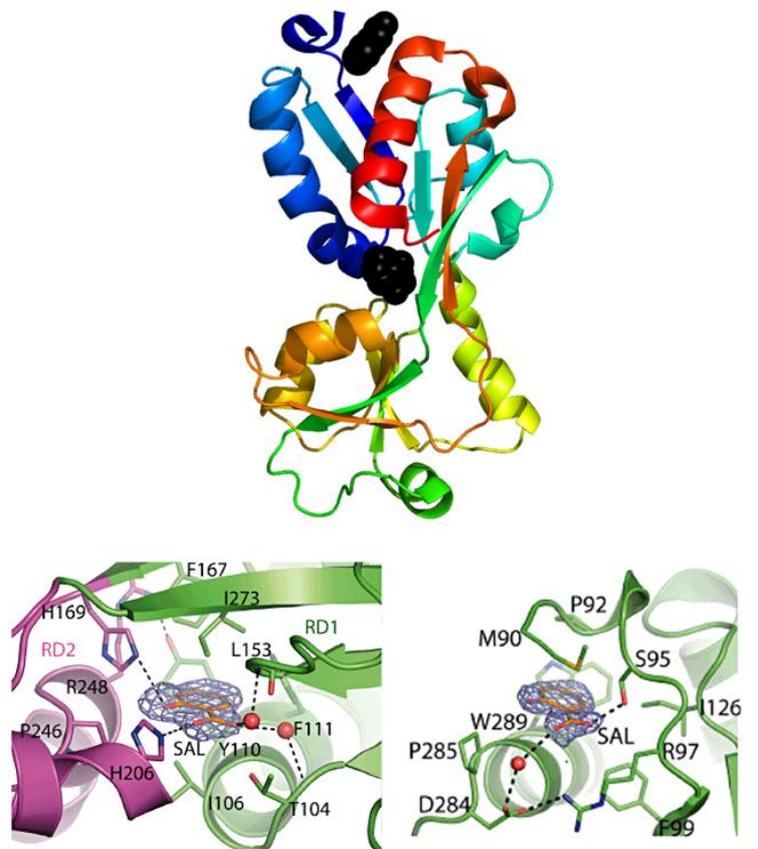


**Figure 1.28: Comparison of the crystal structures of DntR and CbnR homotetramers.** The quarternary structure of full length homotetrameric DntR (left) and CbnR (right). As published by (Smirnova et al., 2004).

### 1.4.1.1 Inducer-bound DntR

While Smirnova and colleagues (Smirnova et al., 2004) presumed that the binding of acetate or thiocyanate in the DntR IBC mimicked that of the natural inducer molecule, in order to further investigate the inducer binding properties of DntR a truncated DntR construct,  $\Delta$ N90DntR, consisting of the IBC only, was produced and crystallized in complex with the DntR effector molecule salicylate (Devesse et al., 2011). Truncating DntR at residue 90 removes the linker helix and N-terminal wHTH domain, and changes quaternary structure from a tetramer to a dimer (Figure 1.18).

The crystal structures obtained (Devesse et al., 2011) revealed two salicylate binding sites (Figure 1.29, top). The first is located in the conserved IBC between RD1 and RD2. In DntR the IBC is lined by hydrophobic residues that accommodate the aromatic ring of the inducer molecule, the hydroxyl and carboxyl groups of which form direct hydrogen bonds with N $\epsilon$ 2 of both H169 and H206 (Figure 1.29 Bottom, left).



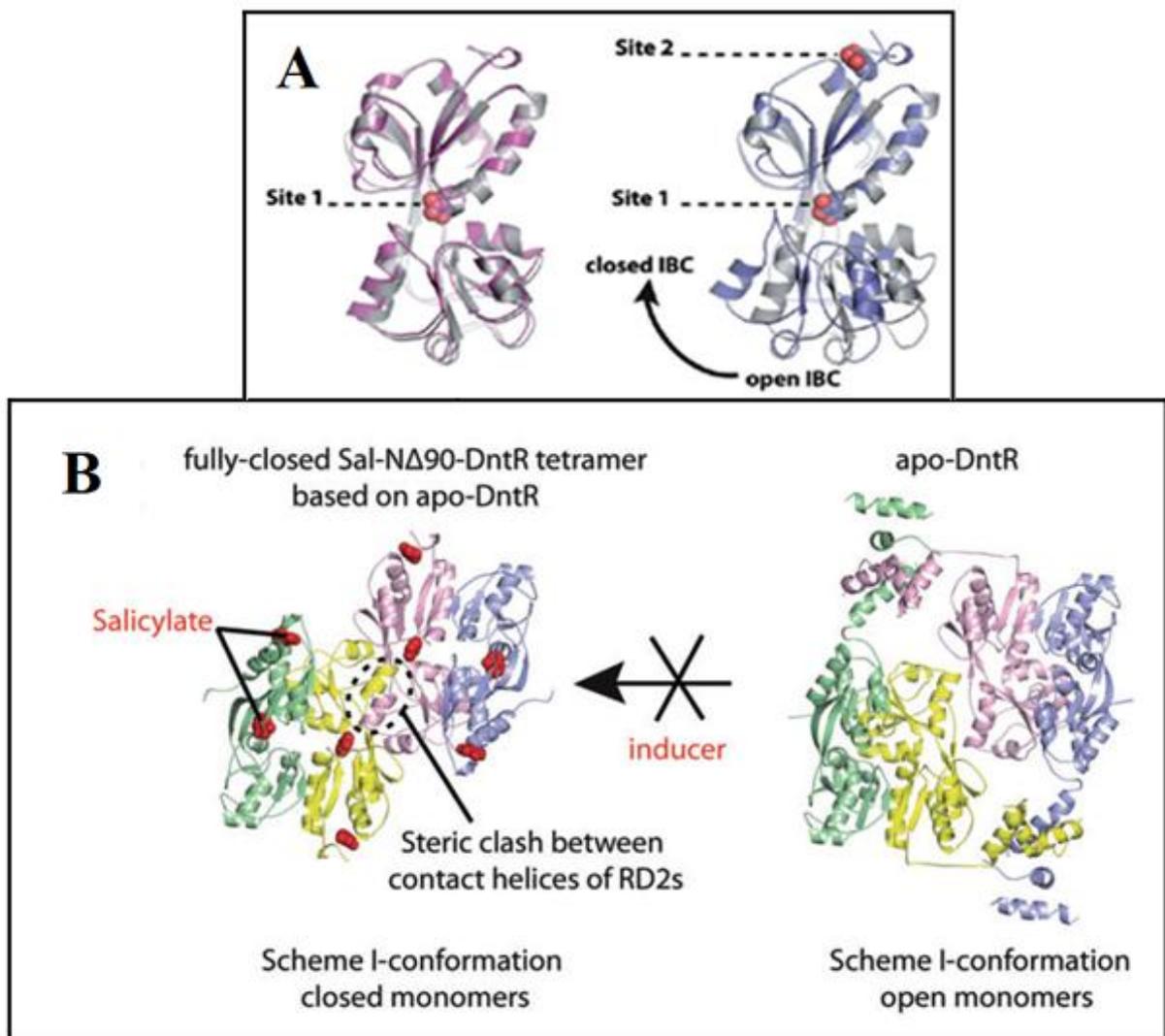
**Figure 1.29: The crystal structure of salicylate-bound  $\Delta$ N90-DntR (PDB 2Y7K).** Top: The  $\Delta$ N90-DntR monomer with two salicylate molecules bound (shown in black spheres). Bottom: Salicylate-bound in the conserved DntR IBC (left) and in the vicinity of secondary binding site (right). Figure from (Devesse et al., 2011).

Unexpectedly salicylate was also found to bind in a second site formed by residues close to the N- and C-termini of the  $\Delta N90$ DntR construct (Figure 1.29, bottom right). Here binding of the aromatic ring is also stabilized by hydrophobic interactions and the salicylate molecule also forms a direct hydrogen bond with S95 and is involved in a solvent mediated interaction with the side chain of D284.

The existence of a secondary binding site in full length DntR *in vivo* is questionable as its location in the  $\Delta N90$ -DntR is occupied by F90 in the crystal structure of the full length protein. ITC studies (Devesse et al., 2011) on both  $\Delta N90$ DntR and full-length DntR constructs neither confirmed nor rejected the idea of a secondary binding site as data could be successfully fitted against both one and two binding site models (Devesse et al., 2011). However, Devesse and colleagues also noted (Figure 1.30A) that the  $\Delta N90$ DntR construct undergoes a significant conformational change only when salicylate is bound in both of the salicylate binding sites observed. The resulting conformation of the DntR IBD is very similar to that observed in the crystal structure of full length inducer-bound TsaR. Moreover, Devesse and Co-workers also noted that while the structures of *apo*- $\Delta N90$ DntR and of  $\Delta N90$ DntR with salicylate bound only in the IBC were consistent with the formation of compact form DntR homotetramers, the conformation of  $\Delta N90$ DntR with two salicylate molecules was not (Figure 1.30B). They therefore speculated (Figure 1.30 B), as suggested by (Monferrer et al., 2010) based on their crystal structure analysis of TsaR, that the active form of DntR, and thus all homotetrameric LTTRs, was the open conformation as seen in the crystal structure of inducer-bound TsaR and of AphB (see Figure 1.21, bottom)

## 1.5 The activation of LTTRs

The details of the activation mechanism of LTTRs currently remains unknown, although there is a consensus that this involves a conformational change upon binding of ligands in the IBC located between the RD1 and RD2 subdomains of the LTTR IBD (Momany and Neidle, 2012; Pérez-Martín et al., 1994; Schell, 1993). This conformational change is proposed to result in a structural rearrangement of the LTTR DBDs bound to the DNA promoter region causing relaxation of the DNA bend and allowing transcription to occur (see Figure 1.8) (Fisher and Long, 1993; Hryniewicz and Kredich, 1995; van Keulen et al., 1998; Parsek et al., 1995; Piñeiro et al., 1997; Toledano et al., 1994; Wang et al., 1992).



**Figure 1.30: Conformational changes to the DntR IBD upon inducer binding.** (A): The crystal structure of *apo*- $\Delta$ N90DntR (grey) overlaid with the crystal structure of  $\Delta$ N90DntR with one (magenta) and two salicylate binding sites occupied (blue). (B): Illustration showing how, when both salicylate binding sites are occupied, the conformation of DntR IBD, observed in crystal structures is inconsistent with the formation of the compact form homotetramers seen for acetate-/thiocyanate-bound DntR (Smirnova et al., 2004). Figure as published by (Devesse et al., 2011).

However, the usefulness of crystal structures to determine the inactive/active states of full length LTTRs is complicated by the fact that the IBDs in the crystal structures obtained often contain molecules from the crystallisation condition bound in their IBCs. One such example is the crystal structure of full-length DntR (Smirnova et al., 2004) in which the IBCs contain either acetate or thiocyanate. Comparing the crystal structures of ion-bound full length DntR with the structure of the homotetramer of full length CbnR, which was crystallised in the absence of any molecules bound in its IBCs (Muraoka et al., 2003) revealed conformational

differences, which would cause a conformational change of bound DNA. It was therefore argued that the crystal structure of CbnR represented that of an LTTR in an inactive state while that of DntR represented an activated LTTR homotetramer. This conclusion was made in despite that neither acetate nor thiocyanate caused any measureable fluorescence quenching of DntR which would have been indicative of conformational changes as a result of acetate or thiocyanate binding (Smirnova et al., 2004). Nevertheless, the hypothesis that the activated LTTR homotetramers maintained a compact form was supported by the crystal structures of inducer-bound and *apo*-BenM IBD (Ezezika et al., 2007). However, as the crystal structures of the former were produced by soaking crystals of the *apo*-IBD, room for conformational change upon inducer binding was limited by crystal packing.

Recent crystal structure analyses (Devesse et al., 2011; Monferrer et al., 2010) suggest that homotetrameric LTTRs actually undergo a conformational change in which the LTTR changes from a compact conformation into a more open form. This is, however, still only a hypothesis. While several crystal structures show that full length LTTR homotetramers can adopt either compact (*i.e.* CbnR, (Muraoka et al., 2003); DntR, (Smirnova et al., 2004); BenM, (Ruangprasert et al., 2010)) or open conformations (TsaR, (Monferrer et al., 2010); ArgP, (Zhou et al., 2010); AphB, (Taylor et al., 2012); PA01 and PA0477 from *Pseudomonas aeruginosa*, (unpublished, PDB codes 3FZV and 2ESN respectively)), to date no LTTR homotetramer has been observed to adopt both open and compact states, questioning whether or not both conformations are available for a given protein. Moreover, in the single case where the crystal structure of a full length LTTR (TsaR; Monferrer et al., 2010) is available in both *apo*- and inducer-bound forms both homotetramers adopt an open state and no conformational change which might result in the relaxation of the bend of bound DNA is observed as a result of the inducer binding.

## 1.6 Aim of thesis

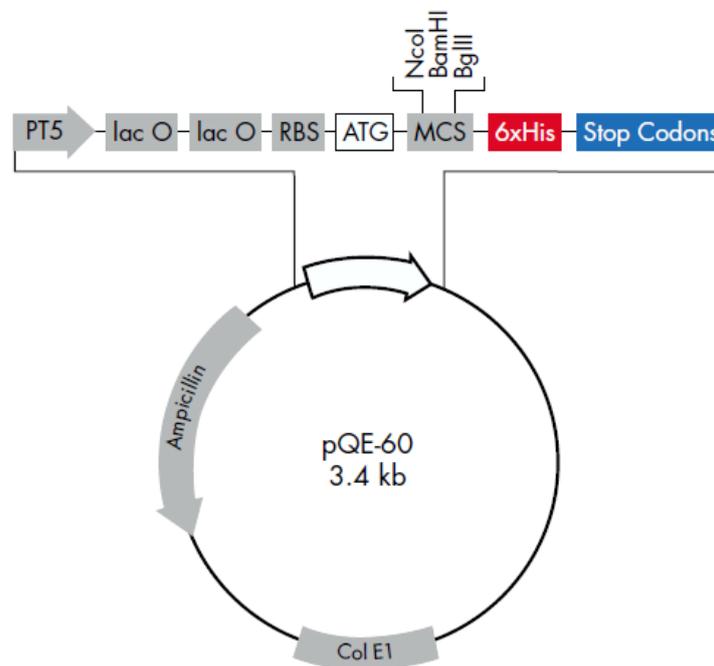
Given the above it is clear that the elucidation of the activation mechanisms of LTTR homotetramers remain elusive and more structural evidence is required. It is also clear that the restrictions imposed on the conformation of LTTRs by crystal packing will need to be overcome by studying the structures of LTTR proteins in solution. The work described in this thesis was directed at shedding more light on the activation mechanism of LTTRs. To this end

the LTR DntR was used as a model system for structural studies employing both X-Ray crystallography and Small Angle X-Ray Scattering (SAXS) aimed at elucidating the structures of *apo*- and inducer-bound DntR and that of a point mutant, H169TDntR, which has previously been shown to show activity in the absence of an inducer molecule (Lönneborg 2012).

## 2. Materials and methods

### 2.1 Expression and purification

DntR DNA on a pQE-60 plasmid (Figure 2.1) was supplied by Rosa Lönneborg, University of Stockholm, Sweden. This expression vector confers ampicillin resistance to the bacteria, adds a C-terminal hexa-histidine tag (His<sub>6</sub>) to the inserted protein sequence and allows for protein overexpression upon the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) via the T5 promoter/lac operator transcription–translation system.



**Figure 2.1: The pQE-60 plasmid.** The plasmid contains a *beta-lactamase* conferring ampicillin resistance and expresses a 6x C-terminal Histidine tag via a *lac* operator system under the control of a T5 promoter.

#### 2.1.1 Wild type DntR (DntR-His<sub>6</sub>)

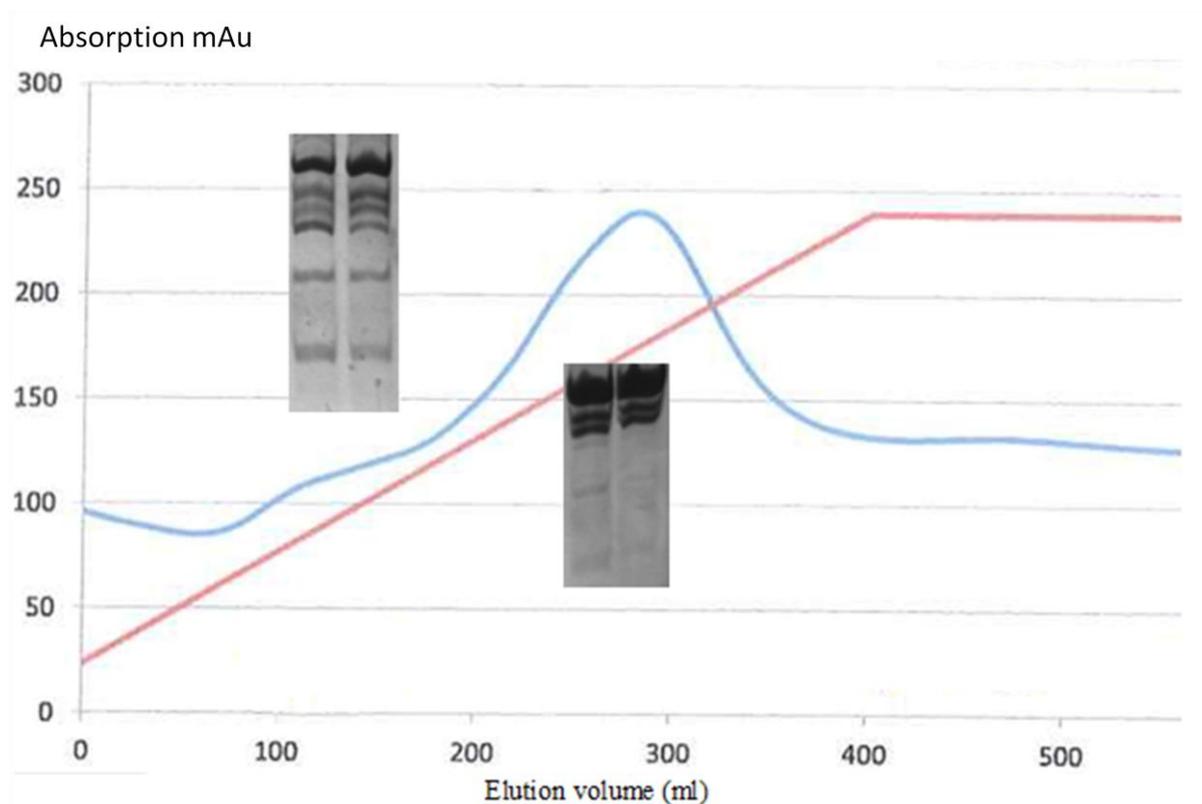
The pQE-60 DntR plasmid was transformed into *E. coli* strain M15[pREP4] (Qiagen). The pREP4 plasmid contained in the genome of these cells constitutively express *lac* repressor causing repression of protein expression prior to IPTG induction and confers kanamycin resistance to the bacteria. 100 µl of M15 cells with 1 µl of plasmid DNA were placed on ice for 30 minutes, heat-shock transformed at 42°C for 90 seconds then put back on ice for 5 minutes. 900 µl of Luria Broth (LB) medium were added to the cell suspension which was subsequently incubated at 37°C with shaking for 1 hour. The suspension was centrifuged at 10,000 x g for a few seconds and approximately 950 µl of the supernatant discarded. The

pellet was resuspended in the remaining 50  $\mu\text{l}$  which was then plated onto LB Agar plates containing 100  $\mu\text{g/ml}$  ampicillin and 50  $\mu\text{g/ml}$  kanamycin.

After overnight growth at 37°C one colony was selected, added to 100 ml of 100  $\mu\text{g/ml}$  ampicillin and 50  $\mu\text{g/ml}$  kanamycin supplemented LB media and grown at 37°C overnight. This pre-culture was then used to inoculate 12 l of LB. Cells were grown at 37°C with shaking in LB media containing 100  $\mu\text{g/ml}$  ampicillin and 50  $\mu\text{g/ml}$  kanamycin for approximately 3 hours until an optical density at  $\lambda=600\text{ nm}$  ( $\text{OD}_{600}$ ) of 0.6 was reached. Protein expression was then induced through the addition of 1 mM (IPTG). To limit degradation, protein expression was tested at different conditions: 37°C for 3, 4 or 5 hours; or overnight at 15, 20 or 25°C. The best ratio of purified protein vs degraded product was obtained for expression at 37°C for 5 hours.

The cells were harvested by centrifugation at 4°C at 5000 x g for 30 minutes. The cell pellet was then resuspended in 50 mM Tris hydrochloride, pH 8.0 containing 300 mM NaCl and 5 mM imidazole (Buffer A) with the addition of 0.1 mg/ml lysozyme and two pellets of cOmplete protease inhibitors (Roche) per liter of culture. The cells were then lysed through ultrasound sonication using 10 bursts of 15 seconds each. The resulting solution was centrifuged at 20,000 x g for 30 minutes to remove cell debris and the supernatant applied to a metal-chelating Hitrap column (GE healthcare Life sciences) previously equilibrated with Buffer A. Upon application of the protein to the column this was washed with buffer A containing 25 mM imidazole to remove non specifically bound proteins. Initial experiments showed that the DntR-His<sub>6</sub> obtained was heavily contaminated with DNA. To remove this in subsequent experiments, the column was washed with buffer A containing 1 M NaCl prior to elution. Elution was then carried out using a gradient of 25 to 500 mM imidazole. Fractions from the column (chromatogram depicted in Figure 2.2) containing a single polypeptide of ~35kDa, as shown by SDS-PAGE analysis (Figure 2.2) were then pooled and dialyzed against 1 l of a buffer comprising 1M NaCl, 25 mM NaH<sub>2</sub>PO<sub>4</sub>-NaOH (pH 8.0) 2 mM MgSO<sub>4</sub>, 1 mM DTT and 17% (v/v) glycerol. The final yield was approximately 1 mg of DntR-His<sub>6</sub> per liter of culture. The protein was concentrated to 5 mg/ml assuming the theoretical extinction coefficient of 18450 M<sup>-1</sup> L<sup>-1</sup>. The concentration was validated by Bradford assays.

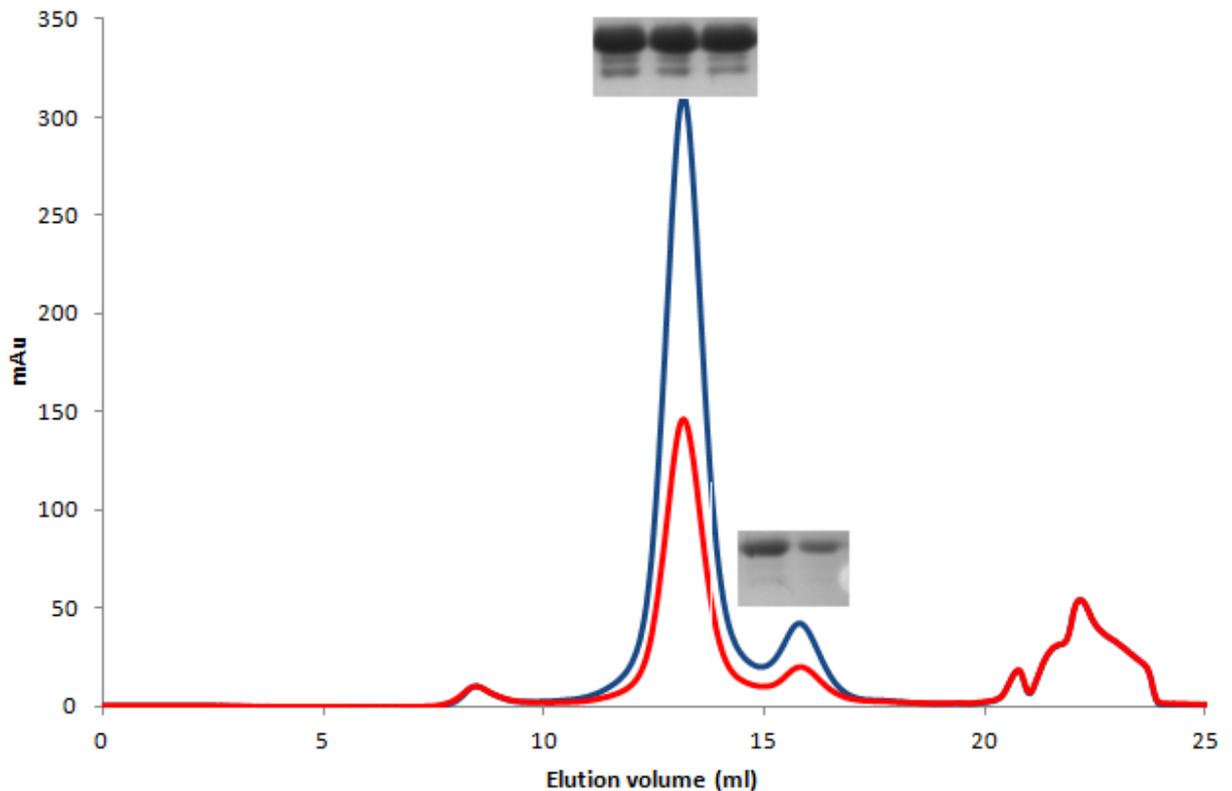
As shown in Figure 2.2, although it was possible to produce relatively pure DntR-His<sub>6</sub> through nickel affinity chromatography the main fractions always contained some degraded protein. To further purify DntR-His<sub>6</sub> from this and other contaminants samples were applied to a gel filtration column.



**Figure 2.2: Purification of DntR-His<sub>6</sub> by nickel affinity chromatography.** Fractions were analyzed using SDS-PAGE. The results corresponding to the shoulder and main peaks are shown as inserts. The absorption at  $\lambda=280$  nm ( $A_{280}$ ) is shown in blue and the imidazole gradient from 0.025-0.5 M is shown in red.

Gel filtration was performed on an Amersham Biosciences Akta Purifier with a Superdex 200 column in a running buffer containing 1 M NaCl, 2 mM MgSO<sub>4</sub>, 1 mM DTT, 5 % (v/v) glycerol, 25 mM NaH<sub>2</sub>PO<sub>4</sub>-NaOH (pH 8.0) at a flow rate of 0.5 ml/min. Fractions of a molecular weight of ~150 kDa, corresponding to a DntR-His<sub>6</sub> tetramer, were pooled dialyzed against 1 M NaCl, 2 mM MgSO<sub>4</sub>, 1 mM DTT, 17% (v/v) glycerol, 25 mM NaH<sub>2</sub>PO<sub>4</sub>-NaOH, concentrated by centrifugation to 5 mg/ml and ultracentrifuged at 50,000 x g in a optima MAX ultracentrifuge to remove remaining aggregated particles. As shown in Figure 2.3 it was never possible to fully separate the DntR-His<sub>6</sub> from degraded products. Attempts to further purify DntR-His<sub>6</sub> were also made using ion exchange chromatography which resulted in no further

improvement (results not shown). As can also be seen in Figure 2.3 DntR exists in equilibrium between tetrameric (~150 kDa) and dimeric states (~75 kDa). As the equilibrium is clearly heavily shifted towards the tetramer this was not addressed further in studies of the tetrameric form of DntR-His<sub>6</sub>.



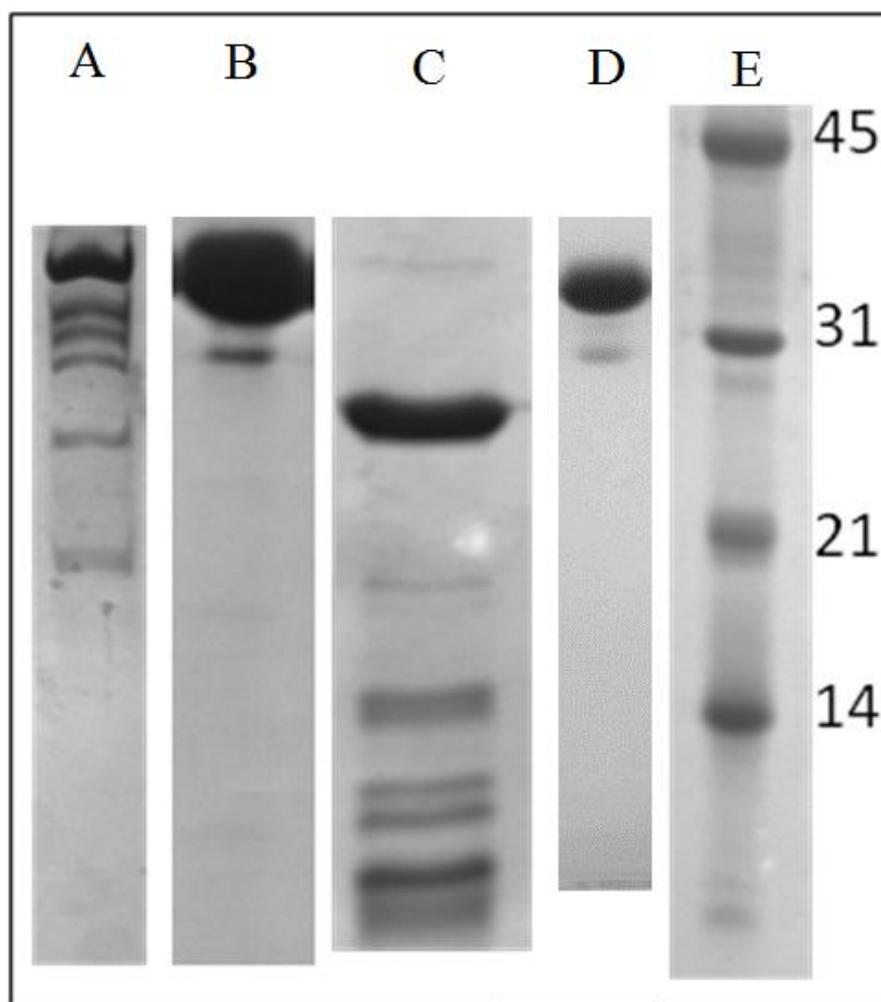
**Figure 2.3: Purification of DntR-His<sub>6</sub> by gel filtration.** Fractions were analyzed using SDS-PAGE and the results corresponding to small and large peaks shown as inserts. A<sub>280</sub> is shown in blue, A<sub>260</sub> in red.

In an attempt to eliminate proteolytic cleavage, expression trials were attempted at different temperatures and over different periods of time (Table 2.1). These trials did not yield significantly different SDS-PAGE expression profiles except for the 3 hour expression at 37°C which had a very poor protein yield (results not shown). Eventually it was found that the addition one c0mplete mini protease inhibitor tablet per liter of culture during cell lysis, reducing total sonication time to 1 minute (15 second pulses, 45 second pauses) and, particularly, carrying out the purification at 4°C and rigorously keeping the cell lysate on ice significantly reduced the amount of degraded protein seen for the full-length constructs

(Figure 2.4). Following this procedure, once purified, no further degradation of the full-length constructs occurred (Figure 2.4).

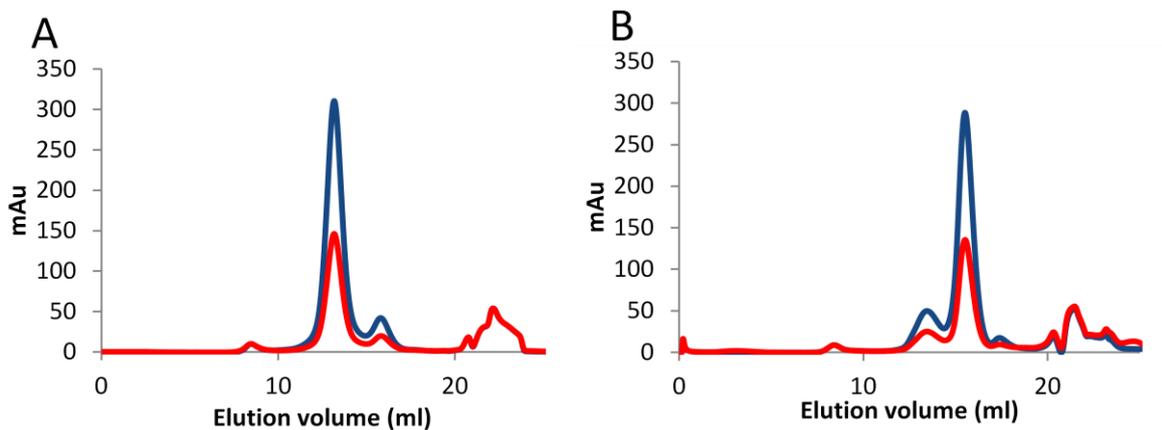
Induction temperature (°C)	37	37	37	25	20	15
Duration of expression (h)	3	4	5	16	16	16

**Table 2.1: Combinations of induction temperatures and hours of expression trialled.** Cells were grown to an OD<sub>600</sub> of ~0.3 at 37°C after which the temperature was changed. Depending on the temperature expression was induced 20-40 minutes later when the OD<sub>600</sub> reached 0.6. The various approaches resulted in no difference in the ratio between degraded and nondegraded full-length DntR as seen in SDS-PAGE analysis.



**Figure 2.4: SDS PAGE analysis of purified DntR-His<sub>6</sub>.** (A): Initial purification trials conducted with low amounts of protease inhibitor and at room temperature. The top band has a mass of ~35 kDa, corresponding to the molecular weight of the DntR-His<sub>6</sub> monomer; several bands corresponding to degraded protein are visible. (B): An example of the result using the current purification protocol involving high amounts of protease inhibitor and carried out at 4°C. (C): Trypsin digests of purified DntR-His<sub>6</sub>. (D): Purified DntR-His<sub>6</sub> left at 4°C for 2 weeks shows no signs of further degradation. (E): Low molecular weight marker.

As expected (Smirnova et al., 2004), DntR elutes on gel filtration as a tetramer. When high amounts of DNase were added to the solution during purification the size exclusion chromatography elution profile was shifted from mainly tetramer to mainly dimer (Figure 2.5). Further experiments are needed in order to verify the relationship between DNA presence and quaternary shape.



**Figure 2.5: Size exclusion chromatography elution profiles of DntR-His<sub>6</sub>.** A<sub>280</sub> shown in blue; A<sub>260</sub> in red. The peak at 12.5 ml elution volume corresponds to a molecular weight of approximately 150 kDa (tetramer) the peak at 16 ml corresponds to 70 kDa (dimer) as determined by calibration measurements (not shown). (A): Standard purification procedure of DntR-His<sub>6</sub>. (B): Purification protocol containing high amounts of DNase.

### 2.1.2 H169TDntR-His<sub>6</sub>

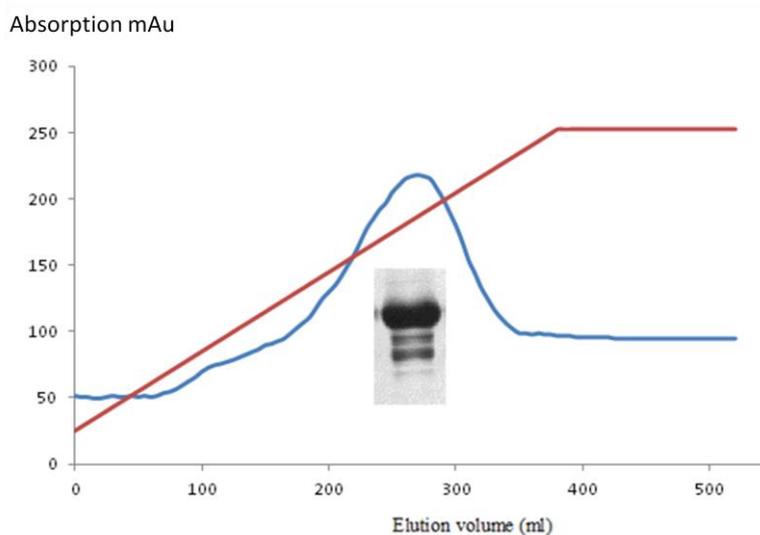
The H169T mutation was introduced into the wildtype pQE60 DntR construct using forward and reverse primers: H169T\_F cggcgcctctttcgaccgcctacgtatgcat; H169T\_R atgcatacgtagcgggtgcgaaagaggcgccg and was amplified by the Polymerase Chain Reaction (PCR) using the procedure shown in Table 2.2.

The PCR products were incubated with 1  $\mu$ l of the restriction endonuclease enzyme DPN1 (New England Biolabs) for 2 hours at 37°C to digest the original methylated template. 50  $\mu$ l of competent *E. coli* strain DH5 $\alpha$  with 5  $\mu$ l of plasmid DNA were then placed on ice for 30 minutes, heat-shock transformed at 42°C for 90 seconds then put back on ice for 5 minutes. 900  $\mu$ l of LB media was added to the cell suspension which was subsequently incubated at 37°C with shaking for 1 hour. The cell suspension was spun down and approximately 950  $\mu$ l of the supernatant discarded. The pellet was resuspended in the remaining 50  $\mu$ l which was plated onto LB agar plates containing 100  $\mu$ g/ml of ampicillin. After overnight growth at 37°C

one colony was selected added to 10 ml of 100 µg/ml ampicillin supplemented media and grown at 37°C overnight. Plasmid DNA was purified through a qiagen miniprep kit and yielded approximately 100 ng DNA / µl. Once the H169T mutation had been validated by DNA sequencing (www.dna.macrogen.com) 1 µl of plasmid DNA was transformed into 100 µl *e. coli* M15[pREP4]. Expression and purification (Figure 2.6) of H169TDntR-His<sub>6</sub> was carried out as described for DntR-His<sub>6</sub>.

Reaction Mixture	PCR Procedure			
50 ng DntR DNA				
1.25 µl H169T_F (10 µM)	Denaturation	30sec	95°C	
1.25 µl H169T_R (10 µM)	Denaturation	10sec	95°C	
2 µl dNTPs (10 mM)	Annealing	10sec	68°C	X30
1 µl pfu polymerase	Elongation	10sec	72°C	
5 µl pfu buffer 5X	Final elongation	5min	72°C	
27 µl Water				

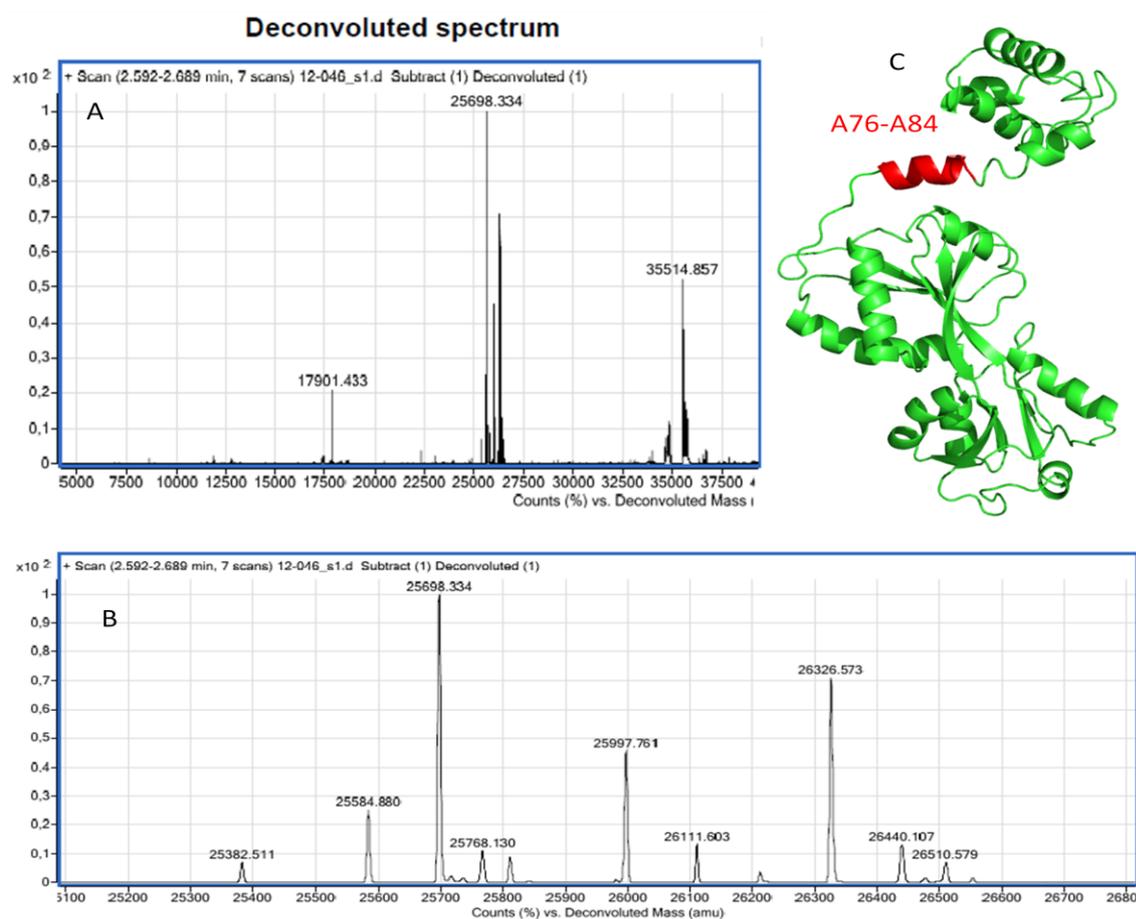
**Table 2.2: The PCR procedure used to create H169TDntR.** The template DNA, primers, nucleotides and polymerase composition of the PCR reaction are shown. Following an initial denaturation step, 30 cycles of denaturation, annealing and elongation followed. Finally the product was elongated for an additional 5 minutes.



**Figure 2.6: Purification of H169TDntR-His<sub>6</sub> by nickel affinity chromatography.** Fractions were analyzed using SDS-PAGE (results shown in insert). A<sub>280</sub> is shown in blue, the imidazole gradient from 0.025-0.5 M is shown in red.

### 2.1.2.1 Mass spectrometry

To validate the H169T mutation and to further investigate the composition of degradation products obtained during purification a sample purified H169TDntR-His<sub>6</sub> was subjected to MALDI-TOF mass spectrometry at the facility at the Institut Biologie Structurale, a platform of the Grenoble Partnership for Structural Biology (<http://www.ibs.fr/platforms/protein-purification/mass-spectrometry>) where samples were analyzed on a 6210 LC-TOF mass spectrometer interfaced with LC pump system (Agilent Technologies). The results are shown in Figure 2.7A. As illustrated, the H169T mutation was confirmed. The analysis also showed that residues at which proteolytic cleavage occurs during protein production are located within the linker helix region connecting the C-terminal IBD with the N-terminal wHTH domain (Figure 2.7B,C).



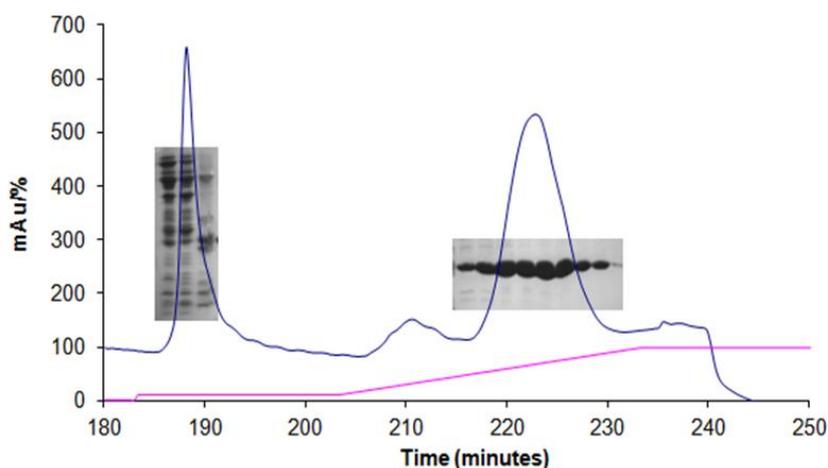
**Figure 2.7: Deconvoluted MALDI-TOF spectrum of H169TDntR-His<sub>6</sub>.** A: full spectrum confirming the H169T mutation (molecular weight of 35515 Da, wild type 35511 Da). B: Different species with molecular weights around 26 kDa. C: The DntR monomer showing the region of DntR-His<sub>6</sub> at which proteolytic cleavage occurs during protein production and purification.

### 2.1.3 $\Delta$ N90DntR and $\Delta$ N90H169TDntR

$\Delta$ N90DntR-His<sub>6</sub> and  $\Delta$ N90H169TDntR-His<sub>6</sub> were produced and purified by a placement student (Marie Christou-Kent, University of Bath, UK) under my supervision. As for full-length DntR, the  $\Delta$ N90DntR DNA on a pQE-60 plasmid was supplied by Rosa Lönneborg, University of Stockholm, Sweden. The H169T mutation was introduced into the  $\Delta$ N90DntR construct using the same primers as for full length H169TDntR and the PCR procedure conducted as shown in Table 2.3. Mutations were validated by DNA sequencing ([www.dna.macrogen.com](http://www.dna.macrogen.com))

Mixture	PCR			
50 ng DntR DNA				
1.25 $\mu$ l H169T_F (10 $\mu$ M)	Denaturation	120sec	95°C	_____
1.25 $\mu$ l H169T_R (10 $\mu$ M)	Denaturation	45sec	95°C	
1 $\mu$ l dNTPs (10 mM)	Annealing	45sec	68°C	X30
1 $\mu$ l pfu pol	Elongation	60sec	72°C	_____
5 $\mu$ l pfu buffer 5X	Final elongation	60sec	72°C	
28 $\mu$ l Water				

**Table 2.3: The PCR procedure used to create  $\Delta$ N90H169TDntR.** The template DNA, primers, nucleotides and polymerase make up the PCR reaction. Following an initial denaturation reaction 30 cycles of denaturation, annealing and elongation followed.



**Figure 2.8: Purification of  $\Delta$ N90DntR-His<sub>6</sub> by nickel affinity chromatography.** Fractions were applied to SDS page with the results corresponding to small and large peak shown as inserts. A<sub>280</sub> is shown in blue and the imidazole gradient from 0.025-0.5 M is shown in pink.  $\Delta$ N90H169TDntR-His<sub>6</sub> showed an elution profile similar to  $\Delta$ N90DntR-His<sub>6</sub> (data not shown). Work of Marie Cristou-Kent (University of Bath, UK).

Transformation into *E. coli* strains DH5 $\alpha$  and M15[pREP4] was conducted as previously described for full length H169TDntR-His<sub>6</sub> (Section 2.1.2). Expression and purification (Figure 2.8) of  $\Delta$ N90DntR- His<sub>6</sub> and  $\Delta$ N90H169TDntR-His<sub>6</sub> was carried out as previously described for  $\Delta$ N90DntR- His<sub>6</sub> (Devesse et al., 2011). The protein yield was much larger than for full length DntR (~7 mg of protein per liter of culture) and since the DNA binding domain was not present in the constructs, there is no issue concerning DNA contamination and with the presence of degradation products.

## 2.2 pDnT

In order to investigate the DNA binding properties of DntR we wished to create a DNA construct (pDnT) containing the sequence of the promoter regions located between the *DntA* and *DntR* genes, in which the DNA binding sites of DntR should reside. The pDnT fragment was constructed with flanking restriction sites for insertion into plasmids to allow amplification in DH5 $\alpha$  cells. The pDnT fragment was constructed (Figure 2.9) using the four primers shown below in PCR reactions.

### primer1

5'-ATGTCGCGCAGATCCATGACCTCACCTCACCTCACCATTATTCATGCTGGTGATTTAACTATCAGACTTGATC-3'

### Primer2

5'-GTTGTCTCCGAATGGCTGCGATTCTAGCGCGTCGATCGGTATAGCGCTATAGATCAAGTCTGATAGTTA-3'

### Primer3 5'-

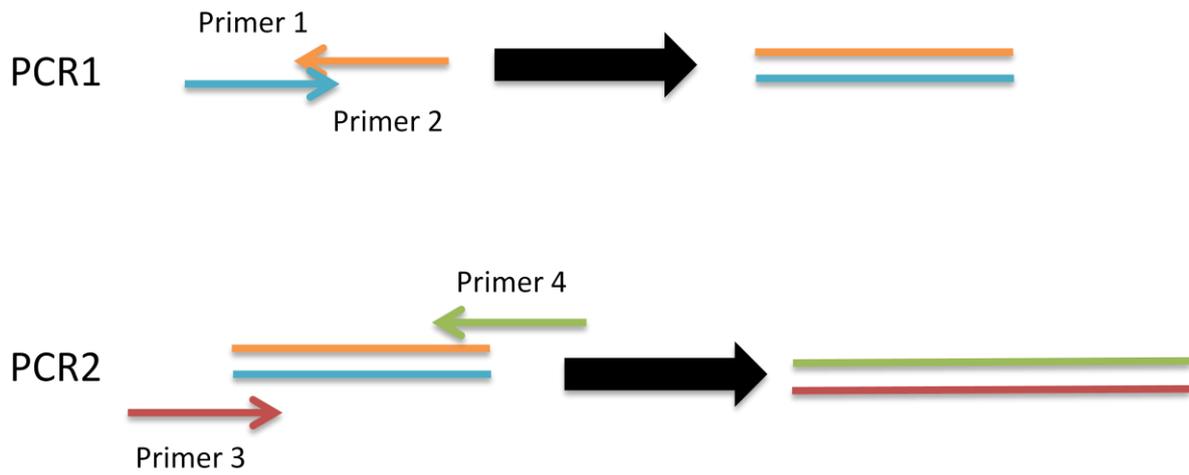
TCGAAGGTAGGCATATGGGTTGAAGACCACCAGCAGATTCAAGTCGATGTCGCGCAGATCCATGACCT-3'

### primer4

5'AAGCTTGAATTCGGATCCCTCGAGTTCATGCAAGCTCTTTTTTTCAGTTGTCTCCGAATGGCTGCGATTCT-3'

PCR reaction 1 was carried out with primer1 and primer2 to create a double stranded DNA fragment containing most of the pDnT sequence. Primers3 and 4 were then added in a second PCR reaction to add restriction sites. Primers 3 and 4 are complementary only to the newly synthesized strands and not primers 1 and 2 and will thus not bind unless the first PCR reaction was successful. The full PCR procedure is schematised in Figure 2.9 and details of the PCR reactions are given in Table 2.4.

**TCGAAGGTAGGCATATGGGTTGAAG**ACCACCAGCAGATTCAAGTCGATGTCGCGCAGATCCATG  
ACCTCACCTCACCTCACCATTATTCATGCTGGTGATTTAACTATCAGACTTGATCATATCGCGATATGG  
CTAGCTGCGCGATCTTAGCGTCGGTAAGCCTCTGTTGACTTTTTTCTCGAACGTACCT**TGAGCTC**  
**CCTAGGCTTAAGTTCGAA**



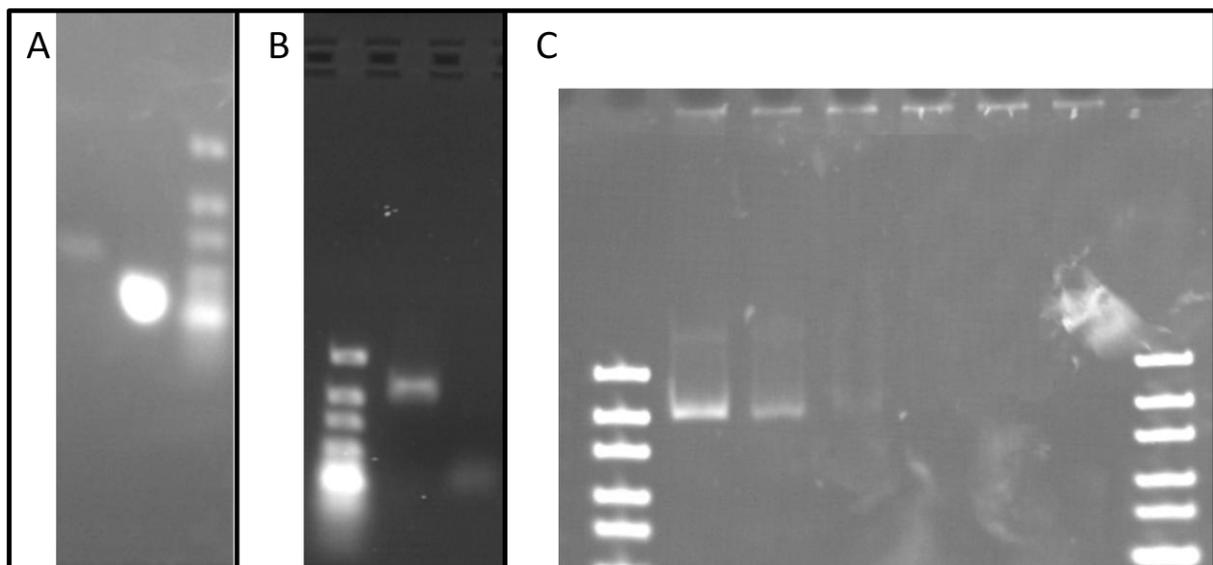
**Figure: 2.9: The PCR reaction for the creation of the pDnT fragment.** Top: sequence of the pDnT fragment, restriction sites shown in red. Bottom: PCR1 consists of two complementary primers of about 70 bp with an overlap of 18 bp. In PCR2 two additional primers are added which complement the newly synthesised strand

Mixture	PCR1			
5 µl Primer1 (2µM)	Denaturation	30sec	98°C	_____
5 µl Primer2 (2 µM)	Denaturation	10sec	98°C	
2 µl dNTPs (10 mM)	Annealing	10sec	45°C	X29
0.5 µl (2U) Phusion pol	Elongation	10sec	72°C	_____
10 µl Phusion buffer 5X	Final elongation	5min	72°C	
27 µl Water				
Mixture	PCR2			
5 µl PCR1	Denaturation	30sec	98°C	_____
5 µl Primer3 (2µM)	Denaturation	10sec	98°C	
5 µl Primer4 (2 µM)	Annealing	10sec	45°C	X29
2 µl dNTPs (10 mM)	Elongation	10sec	72°C	_____
0.5 µl (2U) Phusion pol	Final elongation	5min	72°C	
10 µl Phusion buffer 5X				
27 µl Water				

**Table 2.4: The PCR reactions used to create the pDnT fragment.** The product of PCR1 was used as template DNA for PCR2. For both reactions an initial denaturation step was followed by 29 cycles of denaturation, annealing and elongation and a final elongation step.

The products of PCRs1 and 2 were analyzed on 2% agarose gels in Tris-Borate-EDTA (TBE) buffer comprising 100 mM Tris, 90 mM boric acid, and 1 mM EDTA and stained with CybRSafe (Sigma). As can be seen in Figure 2.10 PCR1, produced a band corresponding to approximately 100 bps while PCR2 produced one corresponding to approximately 170 bps.

The binding between pDnT and DntR-His<sub>6</sub> was tested via electrophoretic mobility shift assays. Native 5% acrylamide gels were prerun at 30 minutes before load of samples comprising DntR in a concentration gradient of 0.16-5  $\mu$ M. However as shown in Figure 2.10C, it was not possible to observe a shift of DNA mobility. As the protein concentration was increased, DNA would no longer enter the gel. To attempt to overcome this, the gels were attempted run either Tris-Borate-EDTA (TBE), Tris pH 8.8 or Tris pH 6.8 with a protein buffer sodium chloride gradient from 100-1000 mM. However, in no case did DNA enter the gel at protein concentrations higher than 1  $\mu$ M and thus a shift of DNA mobility was never observed.



**Figure 2.10: PCR reactions and pDnT DntR binding.** (A): PCR1. Lanes from left to right: product of PCR1, primer1, marker. (B): PCR2: Lanes from left to right: marker, product of PCR reaction 2, primer1. (C): Native PAGE of DntR pDnT binding. The lanes show the mobility of the pDnT fragment after having been incubated with increasing amounts of DntR. From left to right: marker, the addition of 0, 0.16  $\mu$ M, 0.8  $\mu$ M, 1.6  $\mu$ M, 3  $\mu$ M and 5  $\mu$ M DntR.

## 2.3 Crystallization

Prior to all crystallization trials the protein solution was centrifuged at 20,000 x g. All trials carried out were in hanging drops in which 1  $\mu$ l of the protein stock solution ( $\sim$ 5mg/ml) and 1  $\mu$ l of a reservoir solution were equilibrated against 500  $\mu$ l of the reservoir solution. Crystals obtained and used for diffraction data collection were cryoprotected in solutions comprising the relevant reservoir solution supplemented with 20% glycerol then flash frozen in liquid nitrogen.

### 2.3.1 DntR-His<sub>6</sub>

Crystals of DntR-His<sub>6</sub> were grown at 18°C. Crystals of thiocyanate-bound DntR-His<sub>6</sub> were obtained as previously reported by (Smirnova et al., 2004). Crystals of the *apo*-form of DntR-His<sub>6</sub> were obtained using a reservoir solution comprising 0.2 M sodium potassium tartrate, 0.1 M tris (pH 8.0) 5 % (w/v) PEG 6000.

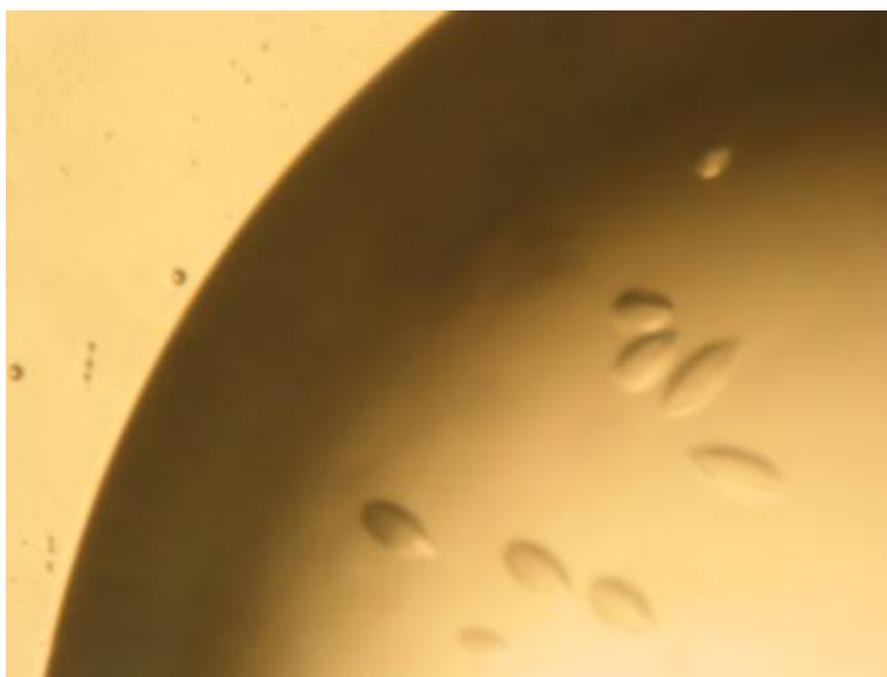
### 2.3.2 H169TDntR-His<sub>6</sub>

Crystals of H169TDntR-His<sub>6</sub> were obtained in both of the conditions described above for DntR-His<sub>6</sub>. However, the crystallization conditions containing no thiocyanate produced crystals diffracting only to  $d_{\min} \sim 12$  Å. Despite several attempts at optimization this diffraction limit could not be improved and the crystal structure of H169TDntR-His<sub>6</sub> was therefore obtained from crystals grown in the presence of thiocyanate.

### 2.3.3 Salicylate-bound DntR-His<sub>6</sub>

To attempt to grow crystals of full length DntR-His<sub>6</sub> bound to salicylate, the protein was incubated with 20, 200 or 5000  $\mu$ M sodium salicylate prior to the setting up of crystallization drops using a reservoir consisting of 0.2 M sodium potassium tartrate, 0.1 M tris (pH 8.0), 5 % (w/v) PEG 6000. The crystals obtained following preincubation with 200  $\mu$ M sodium salicylate are shown in Figure 2.11. No crystal growth was observed in wells containing 5 mM sodium salicylate. Initial tests showed that one of the crystals shown in Figure 2.11 diffracted to  $d_{\min} = 2.15$  Å, the highest resolution obtained thus far for crystals of full length DntR. Full diffraction data collection, structure solution and refinement were carried

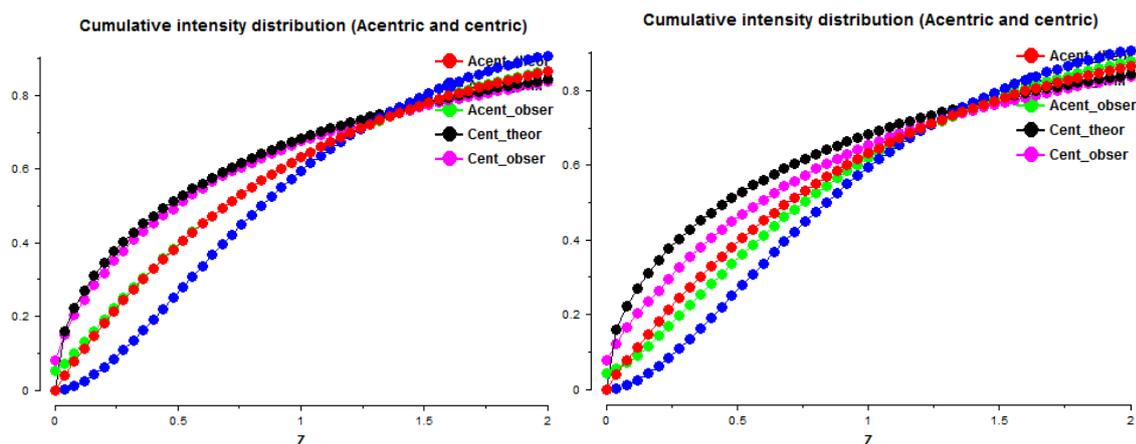
therefore carried out (Table 2.5). At the end of the refinement procedure the electron density for the IBC of molecule A in the asymmetric unit appeared to be consistent with a mixture of two different conformations of the IBC (not shown). The diffraction data were therefore analysed further and it was found that, although this had not been observed for other crystal structures of DntR for which data are reported here, the Cumulative Intensity Distribution indicates partial merohedral twinning (Figure 2.12). A consequence of this is that the crystal symmetry is lower than the  $622$  imposed and that the asymmetric unit contains four rather than two molecules. However, refinement in a lower symmetry spacegroup did not improve the solution (*i.e.* the electron density in the vicinity of the IBC was still disordered). Moreover, refinement in either  $p6_522$  or  $p6_5$  spacegroups did not indicate that salicylate was bound in the IBC. This crystal structure will thus not be discussed further.



**Figure 2.11: Crystals of full-length DntR grown in the presence of salicylate.** DntR was pre-incubated with  $200\ \mu\text{M}$  sodium salicylate. Crystallisation drops comprised  $1\ \mu\text{l}$  pre-incubated DntR and  $1\ \mu\text{l}$  of a precipitant solution containing  $200\ \mu\text{M}$  sodium salicylate (see main text).

Dataset	DntR in presence of Salicylate
Beam-line	ID23-1 (ESRF)
Wavelength	0.976
Space group	P 6 <sub>5</sub> 22
Unit cell dimensions:	
a,b,c (Å)	107.65 107.65 292.48
α,β,γ (°)	90.00 90.00 120.00
Resolution range (Å)	47.16 – 2.15
Number of unique reflections	55226
Multiplicity	4.5 (4.6)
Completeness (%)	99.7 (100)
R <sub>merge</sub>	0.088 (0.905)
<I/σ(I)>	10.3 (2.1)
Wilson B factor Å <sup>2</sup>	47.300
Model quality indicators	
R-factor (%)	22.06
Free R-factor (%)	24.86
RMS deviations:	
Bonds (Å)	0.020
Angles (°)	1.375
Ramachandran analysis (%)	
Favored regions	96.86
generously allowed regions	2.42
Disallowed regions	0.72

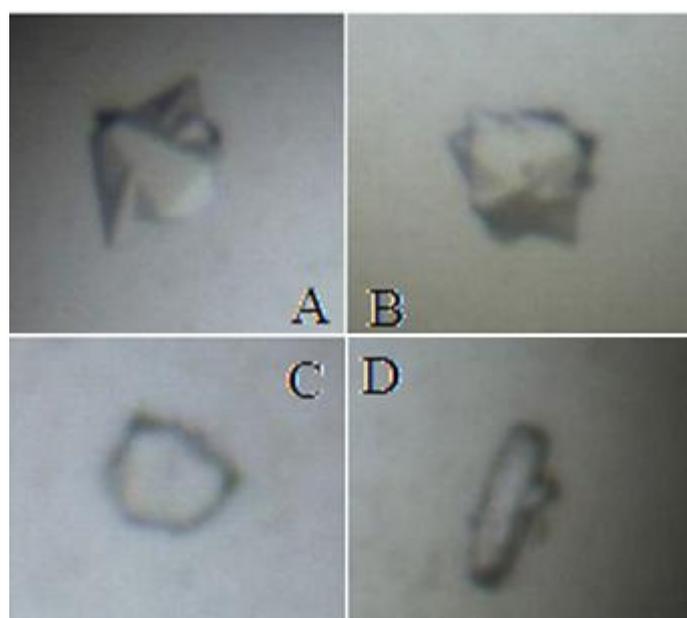
**Table 2.5: Data collection and refinement statistics for crystals of DntR-His<sub>6</sub> grown in the presence of salicylate.** Figures in parentheses are for the highest resolution range for the data set (2.27-2.15 Å).



**Figure 2.12: Cumulative Intensity Distribution plots for crystals of *apo*-DntR-His<sub>6</sub> and DntR-His<sub>6</sub> grown in the presence of salicylate.** *apo*-DntR-His<sub>6</sub> (left) and DntR-His<sub>6</sub> pre-incubated with 200 μM sodium salicylate (right). Theoretical values for centric and acentric reflections (black and red) are shown along with observed values (magenta and green respectively). Theoretical values for a perfectly merohedrally twinned dataset are shown in blue.

### 2.3.3.1 High throughput crystallization trials

In a further attempt to obtain crystals of salicylate-bound DntR-His<sub>6</sub> a stock solution consisting of 5 mg/ml DntR-His<sub>6</sub> in 1 M NaCl, 2 mM MgSO<sub>4</sub>, 1 mM DTT, 17% (v/v) glycerol, 25 mM NaH<sub>2</sub>PO<sub>4</sub>-NaOH, 200 μM sodium salicylate was sent for high throughput screening trials at the EMBL Grenoble High Throughput Crystallization (HTX) lab (<https://embl.fr/htxlab>). In total 576 conditions were screened. After three days crystal-like conformations could be observed in four conditions (Figure 2.13). Attempts to reproduce these in 2 μl hanging drops were unsuccessful.

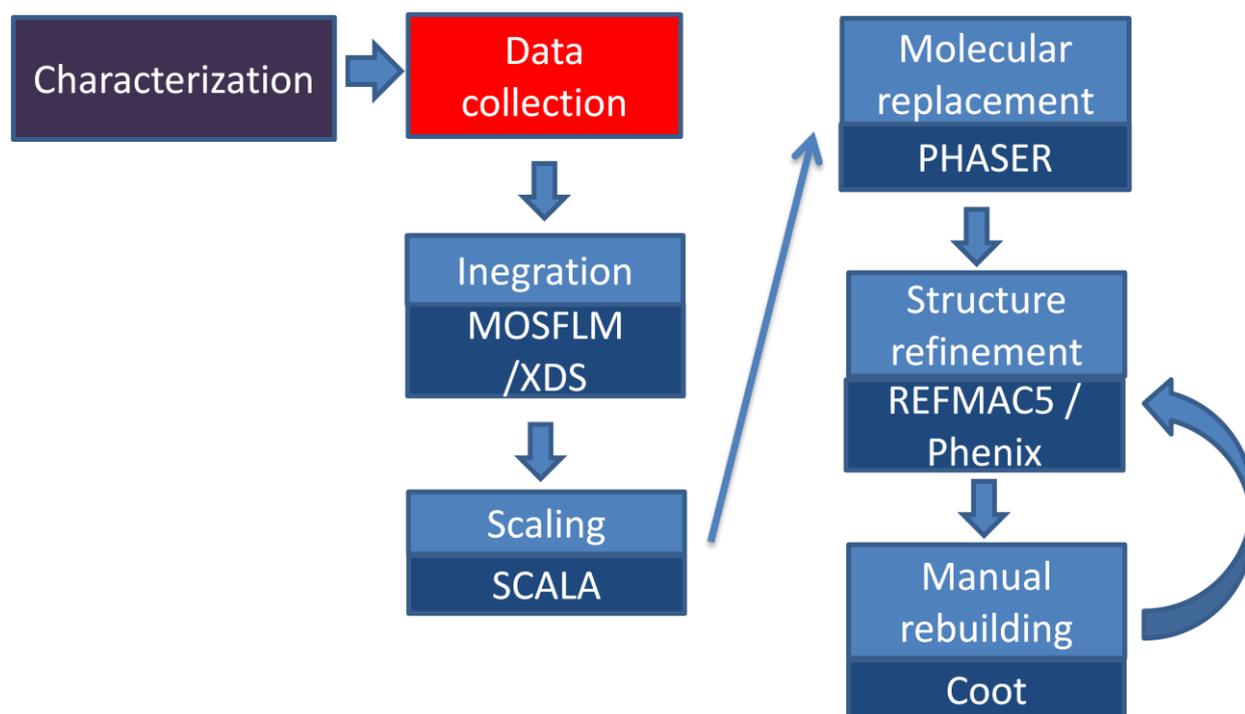


**Figure 2.13: Crystals of potential salicylate-bound DntR-His<sub>6</sub> obtained from high throughput crystallization screening.** The reservoir conditions were as follows: A: 0.2 M Li<sub>2</sub>SO<sub>4</sub>(H<sub>2</sub>O), 0.1 M Tris-HCL pH 8.5, 15% PEG 4000. B: 0.2 M sodium citrate, 20% PEG3350. C: 0.01M ZnSO<sub>4</sub>, 0.1M MES pH 6.5, 25% PEG 550 MME. D: 0.2 M NaCl, 0.1 M tris pH 8.5, 25% PEG 3350.

### 2.3.4 Crystal data collection and refinement

Diffraction data from a single crystal of *apo*-DntR-His<sub>6</sub> were measured on beam-line ID29 (de Sanctis et al., 2012) of the ESRF. Diffraction data from crystals of H169TDntR-His<sub>6</sub> were recorded at ID23-1 (Nurizzo et al., 2006). In both cases, diffraction images were processed using XDS (Kabsch, 2010), intensities merged using the program SCALA (Evans, 2011) and structure factors derived using TRUNCATE (French and Wilson, 1978).

For all crystals analyzed systematic absences in the data sets indicated the space group to be either  $P6_122$  or  $P6_522$ . Structure solution was performed using Molecular Replacement (MR) in the program PHASER (McCoy et al., 2007) using a monomer from the PDB deposition 1UTB (thiocyanate-bound DntR, Smirnova et al., 2004), stripped of solvent and thiocyanate molecules, as a search model. In both cases the MR protocol confirmed the space group to be  $P6_522$ . For *apo*-DntR structure refinement was carried out in REFMAC5 (Murshudov et al., 2011) interspersed with rounds of manual rebuilding in WinCoot (Emsley et al., 2010). Structure refinement of H169TDntR-His<sub>6</sub> was carried out in Phenix (Adams et al., 2010) again interspersed with rounds of manual rebuilding in wincoot. All programs used except XDS and Phenix are distributed via the CCP4 package (Winn et al., 2011). The X-ray data collection, analysis and refinement procedure used throughout this thesis is summarized in Figure 2.14. Crystal characterization and data collection strategies were calculated using EDNA (Incardona et al., 2009) and BEST (Bourenkov and Popov, 2010).



**Figure 2.14: The crystallographic data collection, analysis and refinement strategy used.** Dark blue boxes indicate programs utilized.

### 2.3.5 Thermofluor assays

To assess the thermal stability, thermofluor assays were carried out by the EMBL high throughput crystallisation (HTX) platform (<https://embl.fr/htxlab>). Stock solutions (see section 2.3.3.1 for compositions) containing 5 mg/ml *apo*-DntR, *apo*-H169TDntR and DntR preincubated with 100  $\mu$ M sodium salicylate were sent to the platform for such assays.

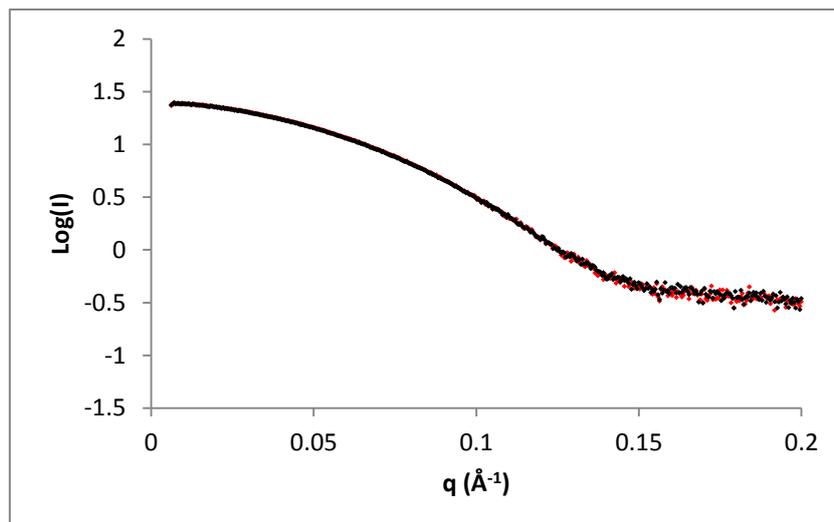
## 2.4 Small Angle X-ray Scattering (SAXS)

All measurements from solutions containing various DntR-His<sub>6</sub> constructs were collected using a Pilatus 1M detector (Dectris Ltd., Baden, Switzerland) in the momentum transfer range  $0.04 < q < 0.61 \text{ \AA}^{-1}$  ( $q = 4\pi \sin \theta/\lambda$ ,  $2\theta$  is the scattering angle between the incident and scattering beam). For static measurements, prior to data collection solutions were centrifuged at 16,000 x g for 10 minutes before being transferred to an automatic liquid handler (Pernot et al., 2013) in which they were stored at 4°C prior to measurements. All measurements of SAXS scattering curves were then conducted at 4°C. Invariant parameters,  $R_g$ ,  $I(0)$ ,  $D_{\max}$  were deduced from the scattering curves as outlined in Section 2.4.3.

### 2.4.1.1 $\Delta N90$ DntR-His<sub>6</sub> and $\Delta N90$ H169TDntR-His<sub>6</sub>

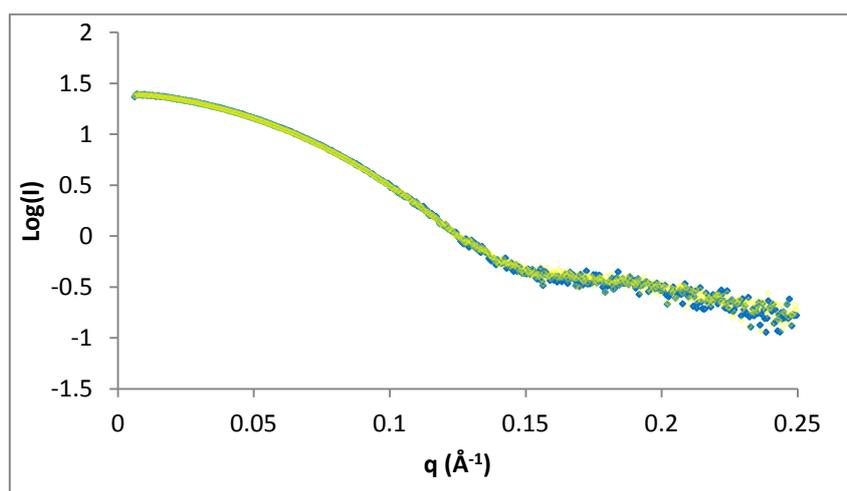
Scattering curves from solutions of  $\Delta N90$ DntR-His<sub>6</sub> and  $\Delta N90$ H169TDntR-His<sub>6</sub> were measured at the ESRF beamline bioSAXS BM29 (Pernot et al., 2013) at  $\lambda=0.992 \text{ \AA}$  at concentrations of 5 and 3 and 1 mg/ml. Samples were flowed through the beam to avoid radiation damage and 10 x 10 frames were recorded for each sample.

Unsurprisingly, the scattering curves of both constructs are very similar (Figure 2.15), an impression confirmed by analysis of the invariant parameters (Table 2.6). A pair-distribution function,  $P(r)$ , was determined from the scattering curve of  $\Delta N90$ DntR-His<sub>6</sub> and a bead model of its solution structure created as outlined in Section 2.4.3. This confirms both  $\Delta N90$ DntR-His<sub>6</sub> and  $\Delta N90$ H169TDntR-His<sub>6</sub> to be dimers in solution.



**Figure 2.15.:** Scattering curves for  $\Delta N90DntR-His_6$  and  $\Delta N90H169TDntR-His_6$ . Scattering curves displaying  $\log(I)$  as a function of  $q$  for solutions of  $\Delta N90DntR-His_6$  (black) and  $\Delta N90H169TDntR-His_6$  (red) after buffer subtraction.

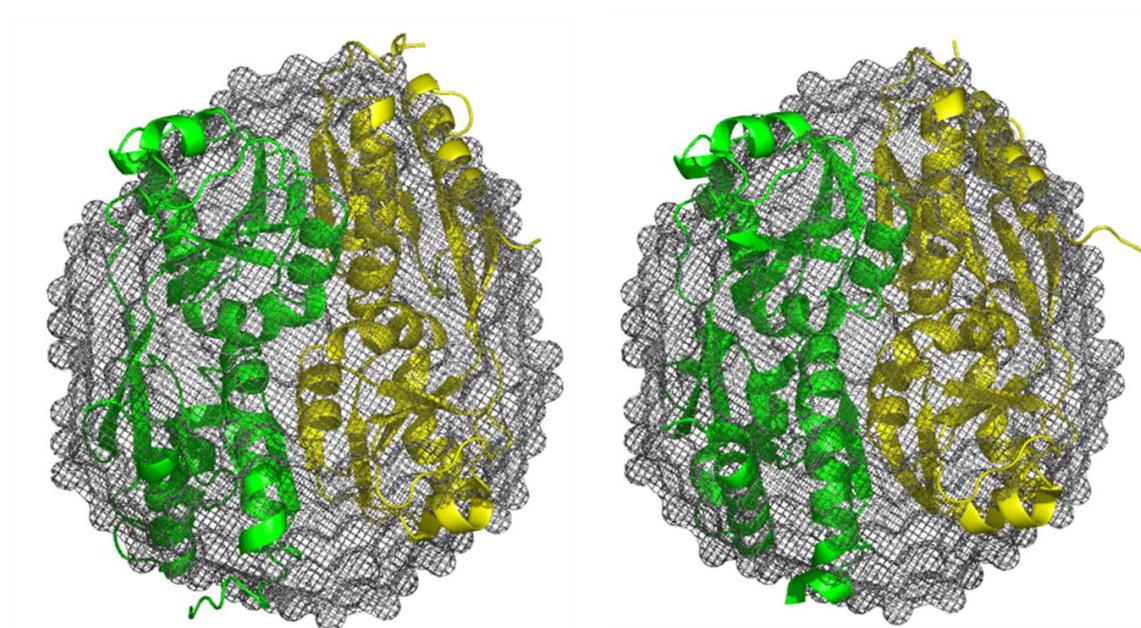
To produce *holo*- $\Delta N90DntR$ ,  $\Delta N90DntR-His_6$  was pre-incubated with 5 mM sodium salicylate. However, this did not induce any significant changes in the scattering curve measured (Figure 2.16). Taken together, Figures 2.15, 2.16, 2.17 and Table 2.6 shows that there is no SAXS measurable conformational change in the structure of the  $\Delta N90DntR-His_6$  dimer upon inducer binding. Any studies using SAXS that aim to help in the elucidation of the active/inactive forms of DntR should thus focus on the full-length protein.



**Figure 2.16:** Scattering curves from solutions of  $\Delta N90 DntR-His_6$  construct in the absence and presence of 5 mM salicylate. Scattering curve displaying  $\log(I)$  as a function of  $q$  for solutions of *apo*- $\Delta N90DntR-His_6$  (green) and  $\Delta N90DntR-His_6$  in the presence of 5 mM salicylate (blue). In both cases the curves displayed are those obtained after buffer subtraction

Invariant parameters	<i>apo</i> - $\Delta$ N90DntR	<i>apo</i> - $\Delta$ N90H169TDntR	<i>holo</i> - $\Delta$ N90DntR
$D_{\max}$ (Å)	65	63	64
$R_g$ (Å)	25.12 +/- 0.43	25.32 +/- 0.33	25.08 +/- 0.38

**Table 2.6:** Invariant parameters obtained from the scattering curves of  $\Delta$ N90DntR constructs.



**Figure 2.17:** *ab initio* model of the solution structures of  $\Delta$ N90DntR-His<sub>6</sub>. The *ab initio* envelope (grey mesh) overlaid with the crystal structures of the *apo*- $\Delta$ N90DntR dimer (left) and *holo*- $\Delta$ N90DntR dimer (right) (Devesse et al., 2011).

#### 2.4.1.2 DntR-His<sub>6</sub>

Scattering curves from solutions of *apo*-DntR-His<sub>6</sub> were measured on the ESRF bioSAXS beamline ID14-3 (Zerrad et al., 2011) at  $\lambda=0.931\text{\AA}$ . 10 frames each of 10 seconds exposure time were recorded for each measurement. Repeated exposure induced no difference in the SAXS scattering curves suggesting that radiation damage had not occurred. Scattering curves from DntR-His<sub>6</sub> in its *apo*-state were measured at concentrations 3, 1 and 0.5 mg/ml. To produce DntR-His<sub>6</sub> in its *holo*-state solutions containing 1 mg/ml DntR were incubated with 5, 20, 100 or 5000  $\mu\text{M}$  sodium salicylate for 10 minutes prior to measurements. For each sample 10 x 10 frames were collected. Sodium salicylate at high concentrations induced

interparticle interactions. For this reason such measurements conducted at a DntR-His<sub>6</sub> concentration of 3 mg/ml, although carried out, had to be discarded.

### 2.4.1.3 H169TDntR-His<sub>6</sub>

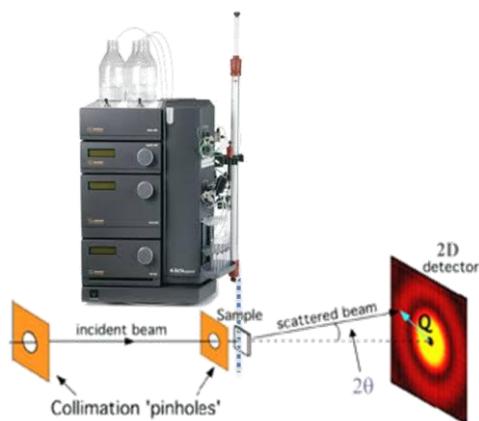
Scattering curves from solutions of *apo*-H169TDntR-His<sub>6</sub> were measured at the ESRF beamline BM29 at  $\lambda=0.992 \text{ \AA}$  at a concentration of 0.7 mg/ml. In order to achieve higher resolution data 10 x 10 frames each of 2 seconds exposure time were recorded and averaged. Samples were flowed through the beam to reduce radiation damage.

## 2.4.2 HPLC-coupled SAXS experiments

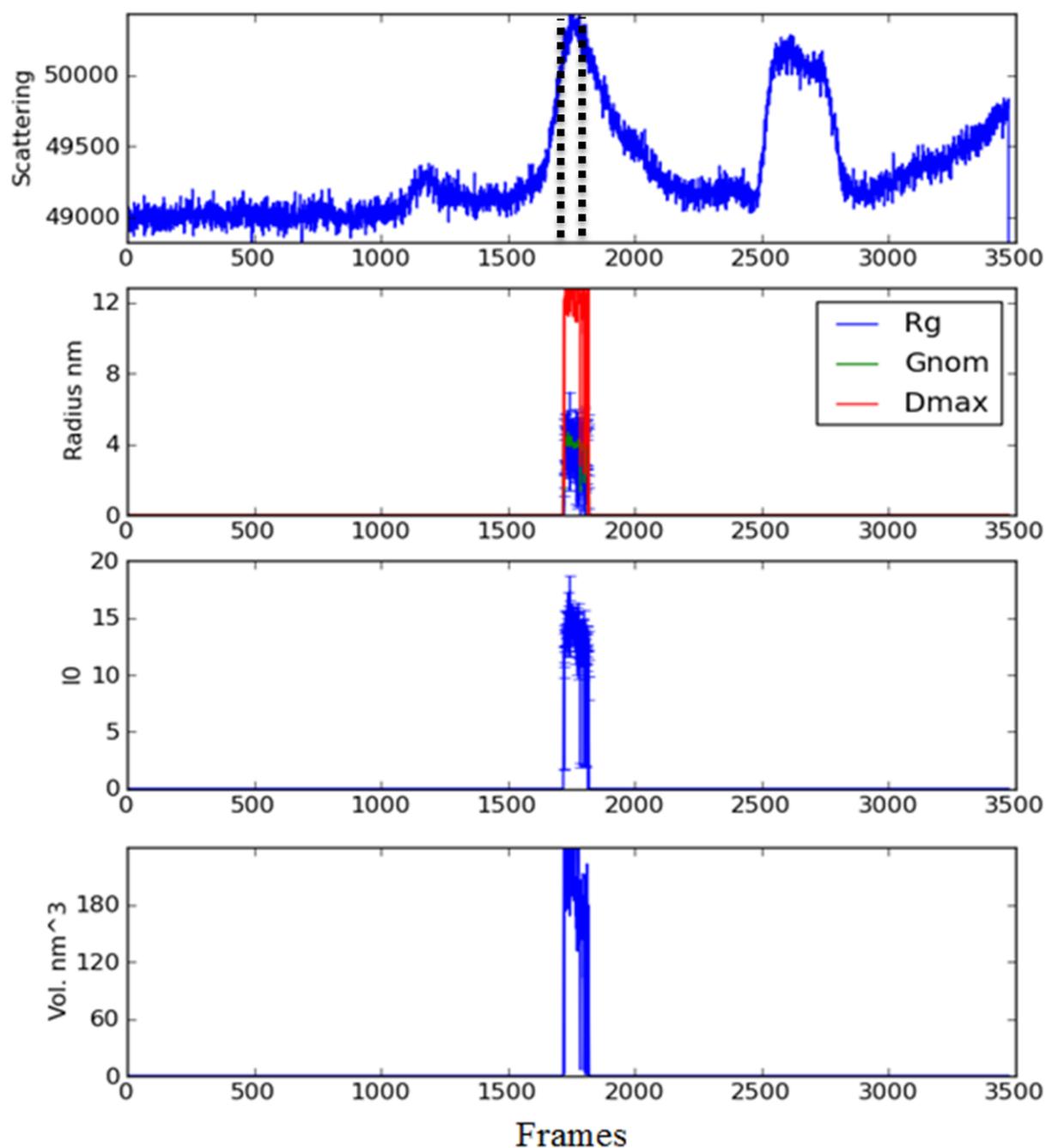
Measurements of the scattering curves of solutions of DntR constructs that were subject to gel-filtration immediately prior to exposure to X-rays were also recorded at the ESRF beamline BM29. The beamline setup, shown in Figure 2.18, makes it possible to load samples on to different kinds of HPLC columns, in this case a gel filtration SD200 column, and to inject the eluent directly into the X-ray beam. This allows for a quasi-instantaneous measurement of the scattering curves of species that would otherwise be impossible to separate due to fast exchanging equilibria. It also makes it possible to separate any aggregated particles. As the flowthrough from the column is exposed constantly by X-rays, effective exposure time and solution concentration is lower than those when performing static measurements. This causes individual measurements of scattering curves to be noisier than for static measurements. However, as the sample will usually elute across ~1-2 minutes the averaging of many individual measurements of scattering curves allows for the higher resolution data than might otherwise be expected.

In this work gel-filtration-coupled SAXS measurements were carried out for solutions of both *apo*-DntR-His<sub>6</sub> and *holo*-DntR-His<sub>6</sub>. In both cases data were collected at  $\lambda=0.992 \text{ \AA}$  and 3500 frames of each 1 second exposure time were recorded as the eluent flowed through the sample cell. Gel filtration of was performed on a Viscotek GPCmax with a Superdex 200 column. 400  $\mu\text{l}$  protein solution at approximately 5 mg/ml was loaded on the column. To obtain data for *holo*-DntR-His<sub>6</sub>, prior to application solutions were incubated with 5 mM sodium salicylate for 30 minutes. In the absence of sodium salicylate the column was run in a

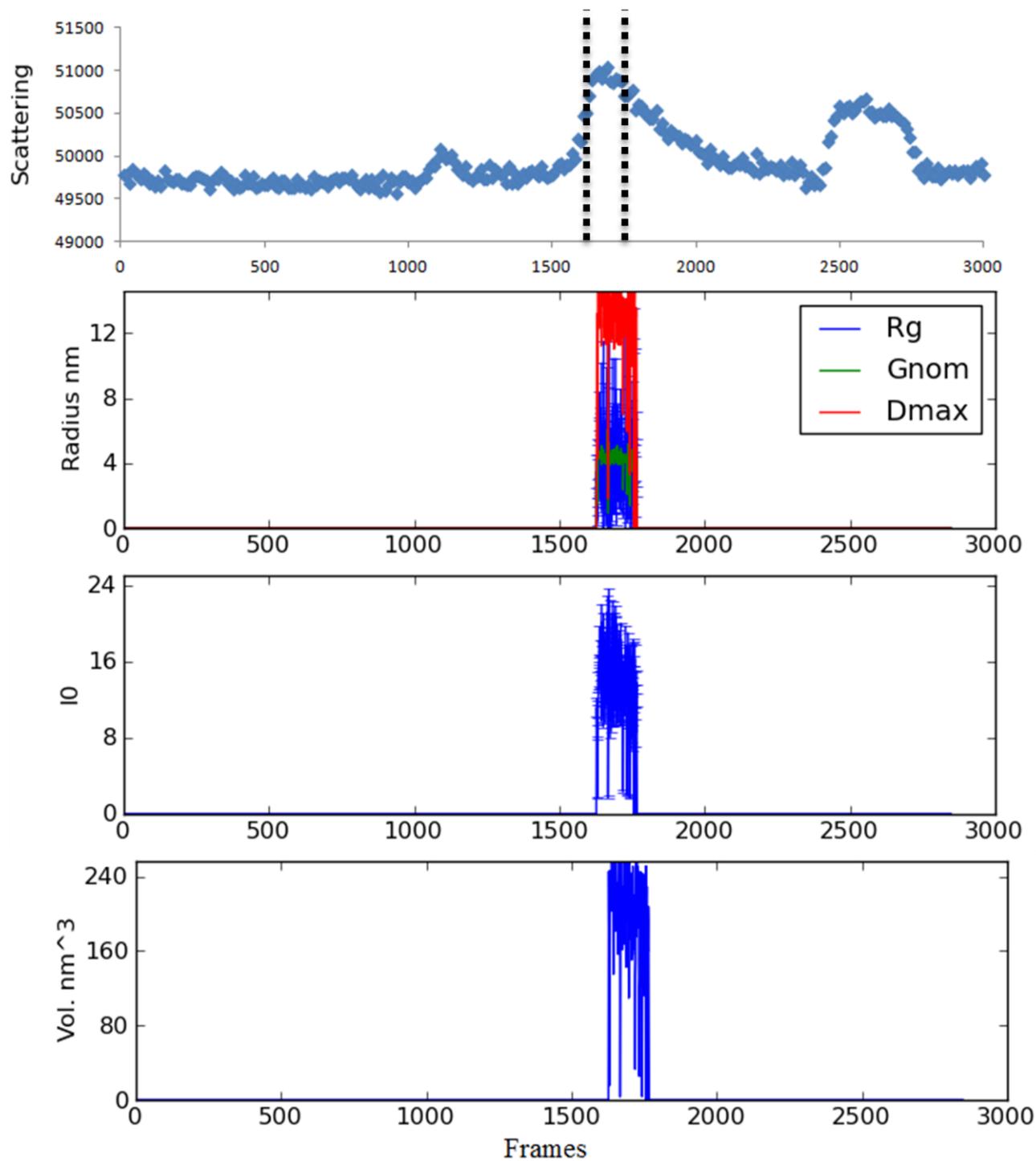
buffer containing 1 M NaCl, 2 mM MgSO<sub>4</sub>, 1 mM DTT, 5 % (v/v) glycerol, 25 mM NaH<sub>2</sub>PO<sub>4</sub>-NaOH (pH 8.0) at a flow rate of 0.5 ml/min. For *holo*-DntR-His<sub>6</sub> 100 μM sodium salicylate was added to the elution buffer. A summary of the results of invariant parameters obtained ( $R_g$ ,  $D_{max}$  and  $I(0)$ ) of the two experiments are shown in Figure 2.19 (*apo*-DntR-His<sub>6</sub>) and 2.20 (*holo*-DntR-His<sub>6</sub>) respectively. As can be seen, in both cases the values of the invariant parameters obtained automatically from the individual scattering curves changes markedly in the last quarter of the frame range presented. These clearly do not contain the same species as the first 75% of the frames and were thus eliminated from subsequent manual processing.



**Figure 2.18: The setup for HPLC-coupled BioSAXS at ESRF beamline BM29.** Top: The experimental hutch of beamline BM29 at the ESRF. The sample environment contains an HPLC setup (lower right corner) that allows for chromatography columns to be mounted. Bottom: schematized diagram of the experimental setup. The eluent from the column runs directly into the sample cell allowing for almost instantaneous measurements of scattering curves. Bottom Figure adapted from [http://www.esrf.eu/UsersAndScience/Experiments/MX/About\\_our\\_beamlines/BM29](http://www.esrf.eu/UsersAndScience/Experiments/MX/About_our_beamlines/BM29).



**Figure 2.19: Summarized output file for HPLC-coupled SAXS experiments of solutions of *apo*-DntR-His<sub>6</sub>.** The top panel shows the total scattering for each frame. The remaining panels show parameters ( $R_g$ ,  $I(0)$ ,  $D_{max}$ ) and porod volume obtained by automated Guinier fits and GNOM plots. The green line corresponds to the quality estimate calculated by AUTOGNOM. Only frames where similar values of  $R_g$  and  $I(0)$  indicated a single species (1715-1809) were used in subsequent processing as indicated by dotted black lines.



**Figure 2.20: Summarized output file for HPLC-coupled SAXS experiments of solutions of *holo*-DntR-His<sub>6</sub>.** The top panel shows the total scattering for each frame. The remaining panels show parameters ( $R_g$ ,  $I(0)$ ,  $D_{max}$ ) and Porod volume obtained by automated Guinier fits and GNOM plots. The green line corresponds to the quality estimate calculated by AUTOGNOM. Only frames where similar values  $R_g$  and  $I(0)$  indicated a single species (1626-1688) were used in subsequent processing as indicated by dotted black lines.

### 2.4.3 SAXS Data analysis

A flowchart showing the protocol used for SAXS data analysis and processing is shown in Figure 2.21. Data analysis for all SAXS measurements was performed in ATSAS 2.4.2 and 2.4.3 available from the EMBL Hamburg group (<http://www.embl-hamburg.de/bioSAXS/software.html>). Radii of gyration ( $R_g$ ) were determined from Guinier plots using PRIMUS (Konarev et al., 2003). In cases where Guinier regions were chosen manually, ranges were chosen within the valid range (in which  $0.8 > q \cdot R_g > 1.2$ ). Pair distribution functions,  $P(r)$ , and the resulting  $D_{\max}$  values were estimated using the indirect Fourier transformation program GNOM (Konarev et al., 2003). Theoretical scattering based on the crystal structures of full-length DntR (Smirnova et al., 2004) and TsaR (Monferrer et al., 2010) were generated and fitted against measured scattering curves using the program CRY SOL (Svergun et al., 1995). Auto processing of frames recorded after HPLC was conducted using EDNA for BioSAXS (Kieffer et al, unpublished).

For purposes of comparison a homology model, constructed using Swiss-Model (Schwede et al., 2003), of a hypothetical DntR homotetramer adopting the same open form as seen in the crystal structure of TsaR (Monferrer et al., 2010) (Figure 1.21) was created.

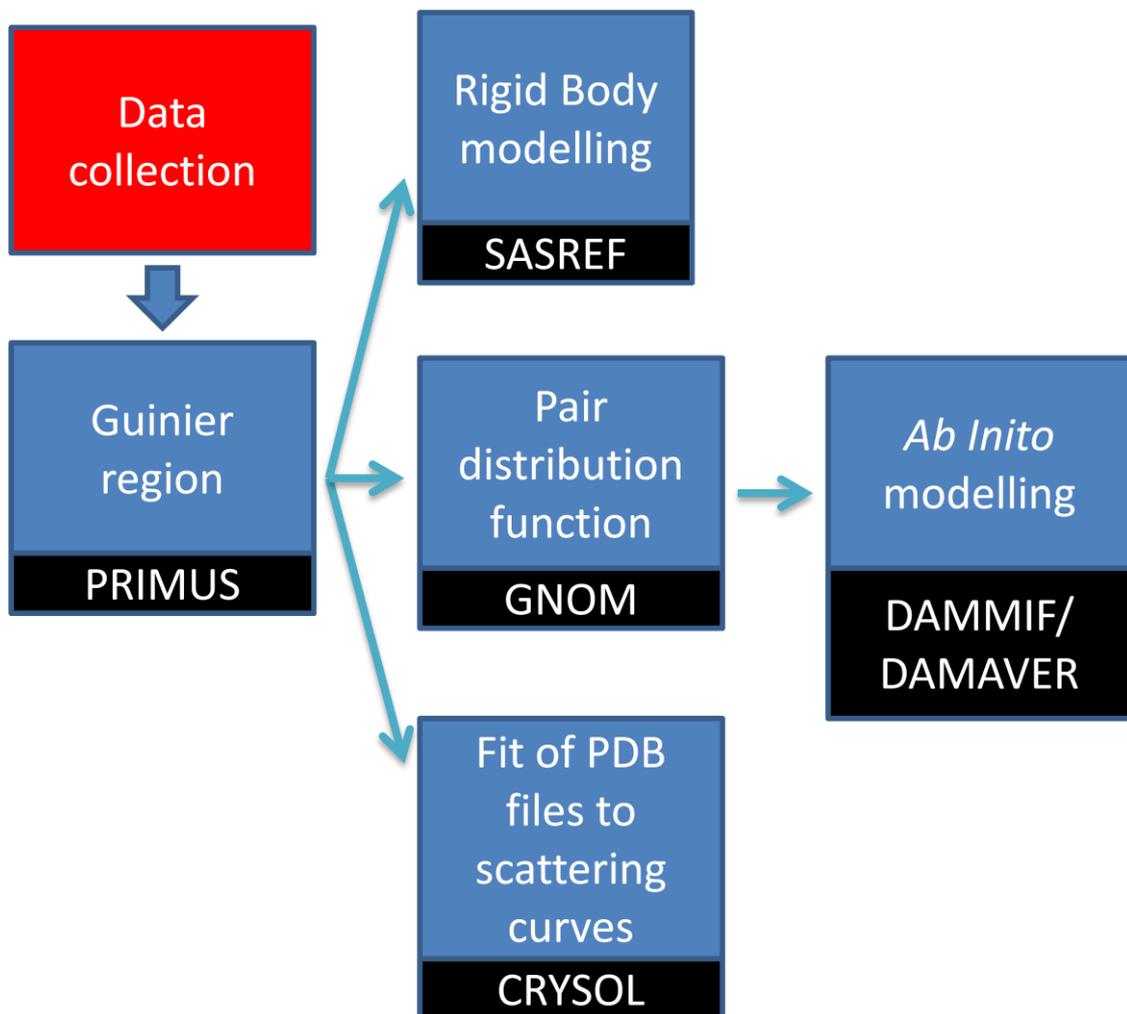
#### 2.4.3.1 *Ab initio* modelling

*Ab initio* models of the solution structures of *apo*-DntR-His<sub>6</sub>, *holo*-DntR-His<sub>6</sub> and *apo*-H169TDntR-His<sub>6</sub> were generated by the same method. Based on generated  $P(r)$  functions as input files, twenty independent bead models, using different randomized seeds, were generated using simulated annealing in the program DAMMIF (Franke and Svergun, 2009). The resulting models were aligned, averaged and filtered using the program package DAMAVER (Volkov and Svergun, 2003). Models were generated both without any symmetry restraints (*apo*-DntR-His<sub>6</sub> and  $\Delta$ N90DntR-His<sub>6</sub>) and with imposed P22 symmetry (*apo*-DntR-His<sub>6</sub>, *holo*-DntR-His<sub>6</sub> and *apo*-H169TDntR-His<sub>6</sub>).

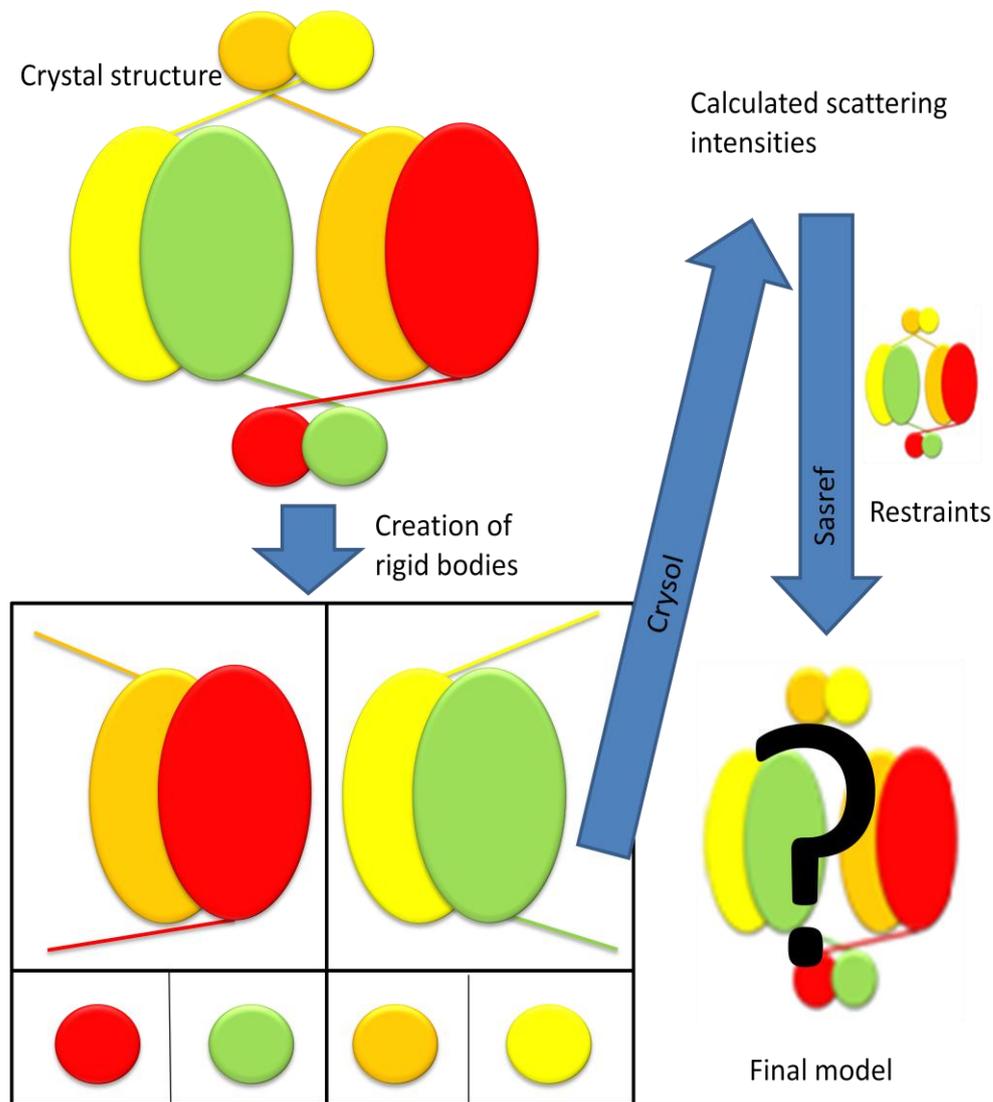
#### 2.4.3.2 Rigid body modelling

In order to obtain improved models, independent of bias introduced through the modelling of  $P(r)$  functions, of the solution structures of *apo*-DntR-His<sub>6</sub> and *holo*-DntR-His<sub>6</sub>, rigid body refinement based on the crystal structure of acetate-bound DntR (Smirnova et al., 2004) was

carried out. Here, dimeric C-terminal units (residues 86-301) and the four N-terminal domains (residues 1-85) were treated as rigid bodies and their positions refined against the experimental scattering curves using SASREF (Petoukhov and Svergun, 2005) with connectivity restraints of 3Å between residue 85 and 86 for each monomer. The SASREF procedure is shown in Figure 2.22.



**Figure 2.21: Flowchart of SAXS data processing.** Programs utilized are shown in black boxes. The determination of Guinier regions is required for any data analysis. It indicates which low angle data points to truncate and is a quality indicator for the dataset. If a Guinier region cannot be determined the data is most likely aggregated or not of sufficient quality for further analysis. The Guinier region determination also provides the invariant parameters  $R_g$  and  $I(0)$  providing information about the size and oligomerisation (*i.e.*, monomer, dimer, tetramer etc) of the molecule in solution.



**Figure 2.22: Rigid body modelling of the solution structures of *apo*- and *holo*-DntR-His<sub>6</sub>.** The domain structure of the (Smirnova et al., 2004) model of the structure of a DntR homotetramer is divided into several rigid bodies. The theoretical scattering intensities of each of the rigid bodies are then calculated in CRYSOLOG. These are then used as input for SASREF along with connectivity restraints ensuring that the main polypeptide chain remains connected. SASREF then refines the position of the rigid bodies within the given restraints against the experimental scattering curve and generates a model of the solution structure.

## 2.5 Figure preparation

All images of crystal structures, SASREF generated rigid body models and *ab initio* envelopes were produced using the molecular program PyMOL ([www.PyMOL.org](http://www.PyMOL.org)) developed by DeLano, WL 2002. Coordinates for previously solved crystal structures were downloaded from the Protein Data Bank ([www.rcsb.org](http://www.rcsb.org)). Electron density maps were exported into a PyMOL readable format using the program FFT from the CCP4 package.

## 3 Results

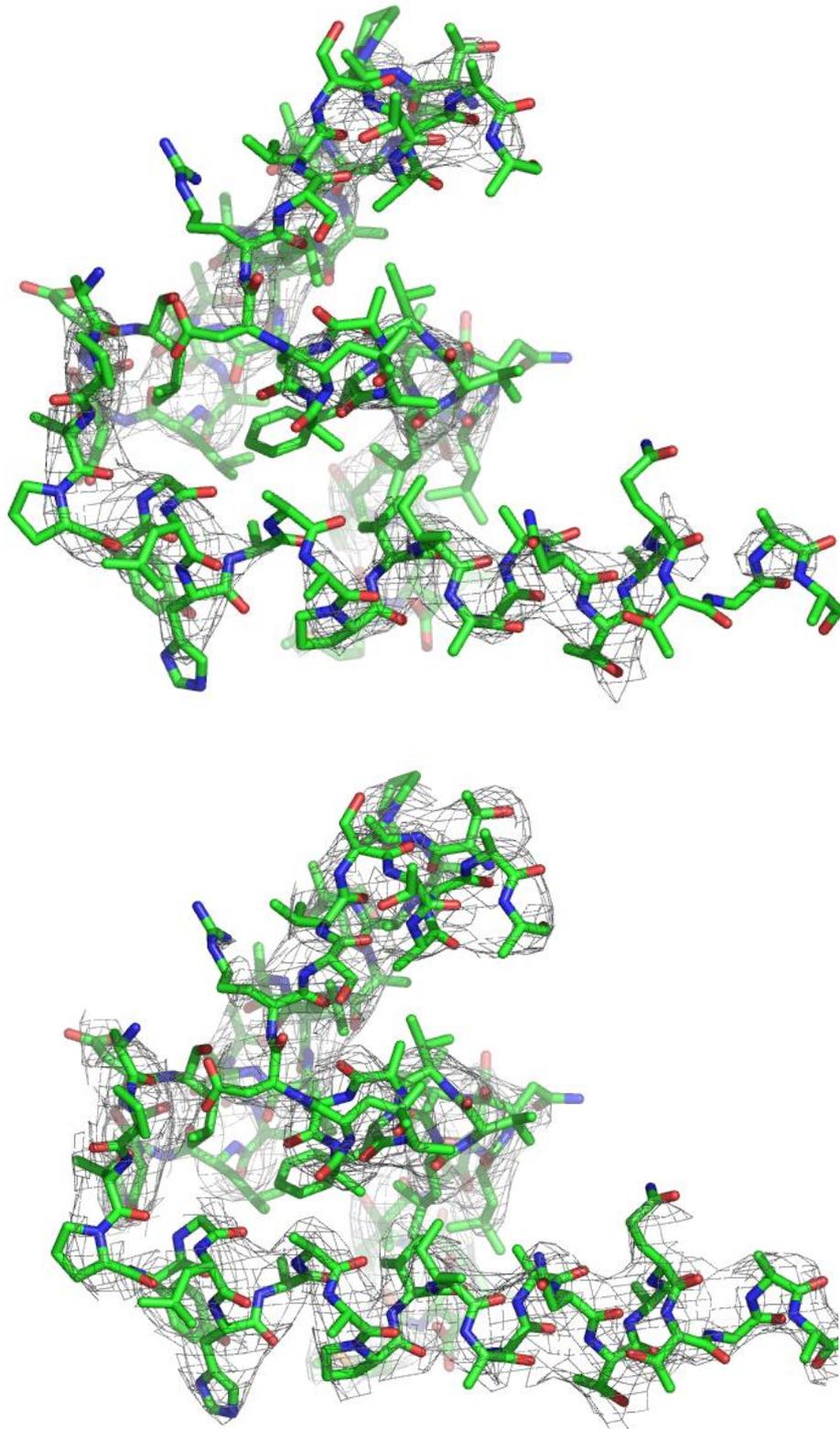
Although a C-terminal His<sub>6</sub> tag was present in all DntR constructs expressed, purified and studied in this thesis work, for simplicity, the –His<sub>6</sub> will be omitted from the construct names in the following Sections (*i.e.* DntR not DntR-His<sub>6</sub> etc).

### 3.1 *apo*-DntR

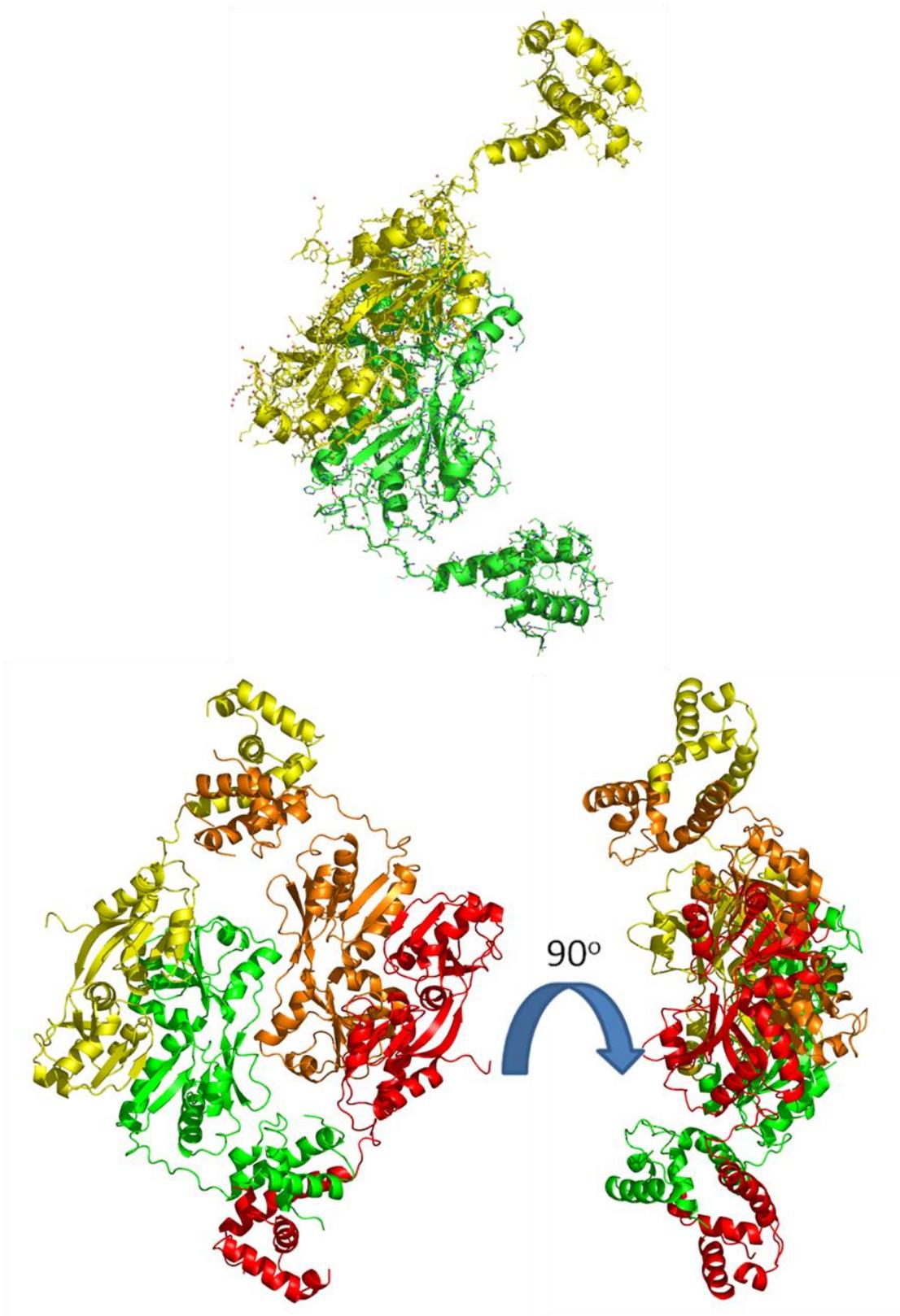
#### 3.1.1 The crystal structure of *apo*-DntR

Based on differences observed between the crystal structure of CbnR (Muraoka et al., 2003) and the model they obtained for the full-length structure of acetate-/thiocyanate-bound DntR, Smirnova and colleagues suggested that the structure they had obtained was that of activated DntR (Smirnova et al., 2004). However as outlined in Section 1.5 the suggestion was made despite the lack of any measureable fluorescence quenching of DntR upon acetate or thiocyanate binding.

In the work presented here, crystals of *apo*-DntR were obtained using the crystallization conditions described in Section 2.3.1 and its crystal structure solved and refined (Table 3.1). This represents the first crystal structure of *apo*-DntR. Moreover, as can be seen in Figure 3.1 during the refinement of the crystal structure it was also possible, although they are still clearly very mobile, to model the linker/HTH domains of the homotetramer to a much larger extent than had been possible in the work described by (Smirnova et al., 2004) and in subsequent crystal structure of DntR point mutants (Lönneborg et al., 2007). The result presented here (Figure 3.2) thus also represents the first complete crystal structure of a full-length DntR.



**Figure 3.1:** Electron density for the N-terminal regions in the crystal structure of *apo-DntR*.  $2mF_o - DF_c$  electron density (mesh), at the end of the refinement procedure for a HTH/linker helix region of a *apo-DntR* monomer contoured at  $1.0 \sigma$  (top) and  $0.5 \sigma$  (bottom). The resulting model of this region is shown in stick representation with carbon atoms shown in green, nitrogen atoms in blue and oxygen atoms in red.



**Figure 3.2: The crystal structure of *apo*-DntR.** Top: The DntR head-to-tail dimer observed in the asymmetric unit for of the *apo*-DntR crystals. Bottom: The crystal structure of the full length *apo*-DntR homotetramer constructed based on symmetry related head-to-tail dimers such that they have coordinates  $(x, y, z)$  and  $(1-x, y, \frac{1}{2}-z)$ .

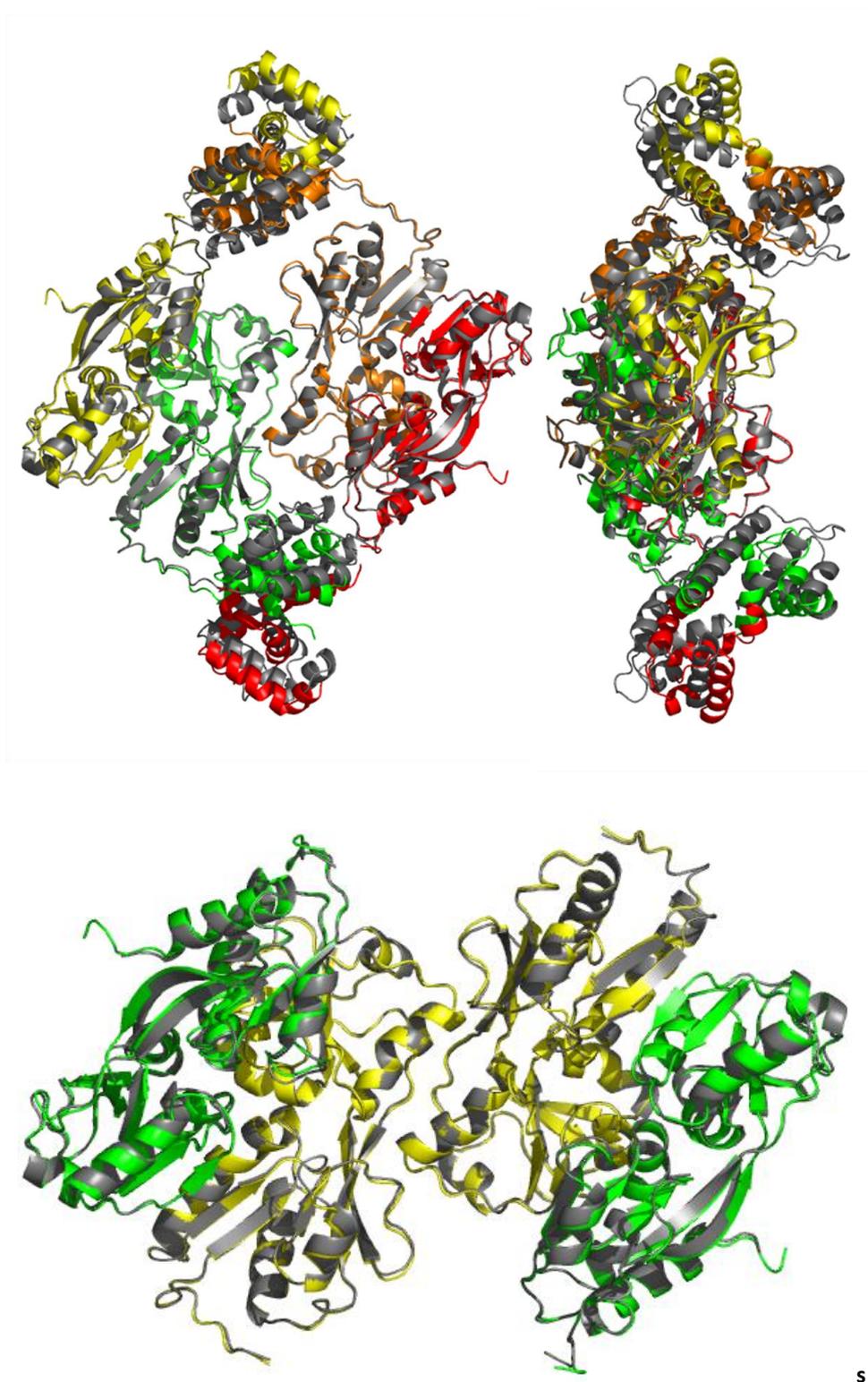
Dataset	<i>apo</i> -DntR
Beam-line	ID29 (ESRF)
Wavelength (Å)	0.976
Space group	P 6 <sub>5</sub> 22
Unit cell dimensions:	
a,b,c (Å)	107.13 107.13 297.84
α,β,γ (°)	90.00 90.00 120.00
Resolution range (Å)	47.14 - 2.64
Number of unique reflections	30064
Multiplicity	7.9 (8.2)
Completeness (%)	99.7 (98.1)
R <sub>merge</sub>	0.102 (1.047)
<I/σ(I)>	13.5 (2.1)
Wilson B factor (Å <sup>2</sup> )	77.547
Refined model composition	
Molecule A	M1-K29, T31-L52, E61-H303
Molecule B	D5-52, E61-E300
Glycerol molecules	9
Water molecules	83
Model quality indicators	
R-factor (%)	19.89
Free R-factor (%)	22.68
RMS deviations:	
Bonds (Å)	0.007
Angles (°)	0.972
B factors (Å <sup>2</sup> )	
<B>	93.8
<B> IBD atoms only	70.6
<B> water molecules	61.7
Ramachandran analysis (%)	
Favored regions	95.2
Generously allowed regions	4.2
Disallowed regions	0.5

**Table 3.1: Data collection and refinement statistics for the crystal structure of *apo*-DntR.** The numbers in

parentheses are for the highest resolution range for the data set (2.79-2.64 Å).  $R = \sum_{hkl} \frac{|F_{hkl}^{obs} - F_{hkl}^{calc}|}{\sum_{hkl} F_{hkl}^{obs}}$

$$R_{merge} = \frac{\sum_{hkl} \sum_j |I_{hkl} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j \langle I_{hkl,j} \rangle}$$

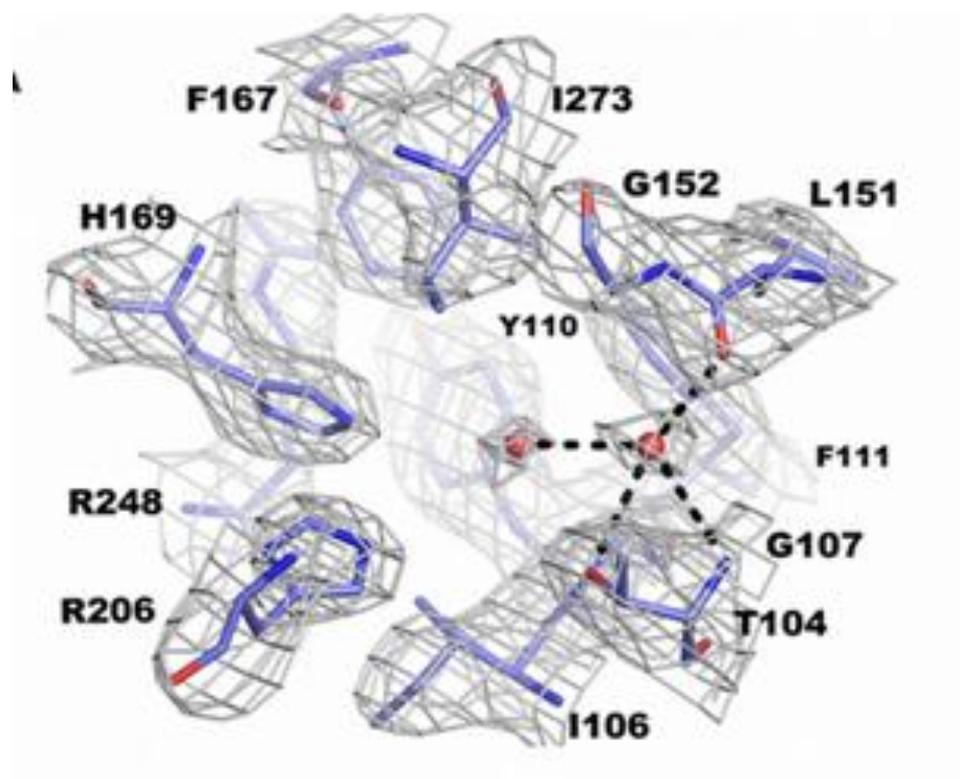
As can be seen in Table 3.1 the crystals of *apo*-DntR obtained appear isomorphous to those obtained for DntR in complex with either acetate or thiocyanate (Smirnova et al., 2004). This is reflected in the crystal structure obtained (Figure 3.2) which, globally, is very similar to that observed for acetate-/thiocyanate-bound DntR. As for the crystal structures of acetate-/thiocyanate-bound DntR the asymmetric unit of the crystal structure of *apo*-DntR contains a head-to-tail dimer with a compact DntR homotetramer being constructed by the association of two symmetry related dimers. Moreover, a superposition of the model of thiocyanate-bound DntR obtained by (Smirnova et al., 2004) and the *apo*-DntR crystal structure elucidated here shows a r.m.s. deviation in  $C_{\alpha}$  positions of 1.04 Å. 1190 residues aligned (Figure 3.3, Bottom) and a r.m.s. deviation of 0.6 Å for the superposition of atoms in the IBD tetrameric core only. Indeed, the only significant difference between the structure reported here and those obtained by (Smirnova et al., 2004) is that the DntR Inducer Binding Cavity (IBC) is devoid ions and contains only ordered water molecules (Figure 3.4). This confirms the truly *apo*-nature of the structure described here. As can be seen from Figure 3.3, when comparing the structures of *apo*- and acetate-/thiocyanate-bound DntR homotetramers there are slight differences in the positions of the WHTH dimers flanking the cores of the IBD tetramers, however, as the electron density was not sufficient to fully model these domains in the crystal structures of thiocyanate- or acetate-bound DntR, it is uncertain whether this difference is an artefact.



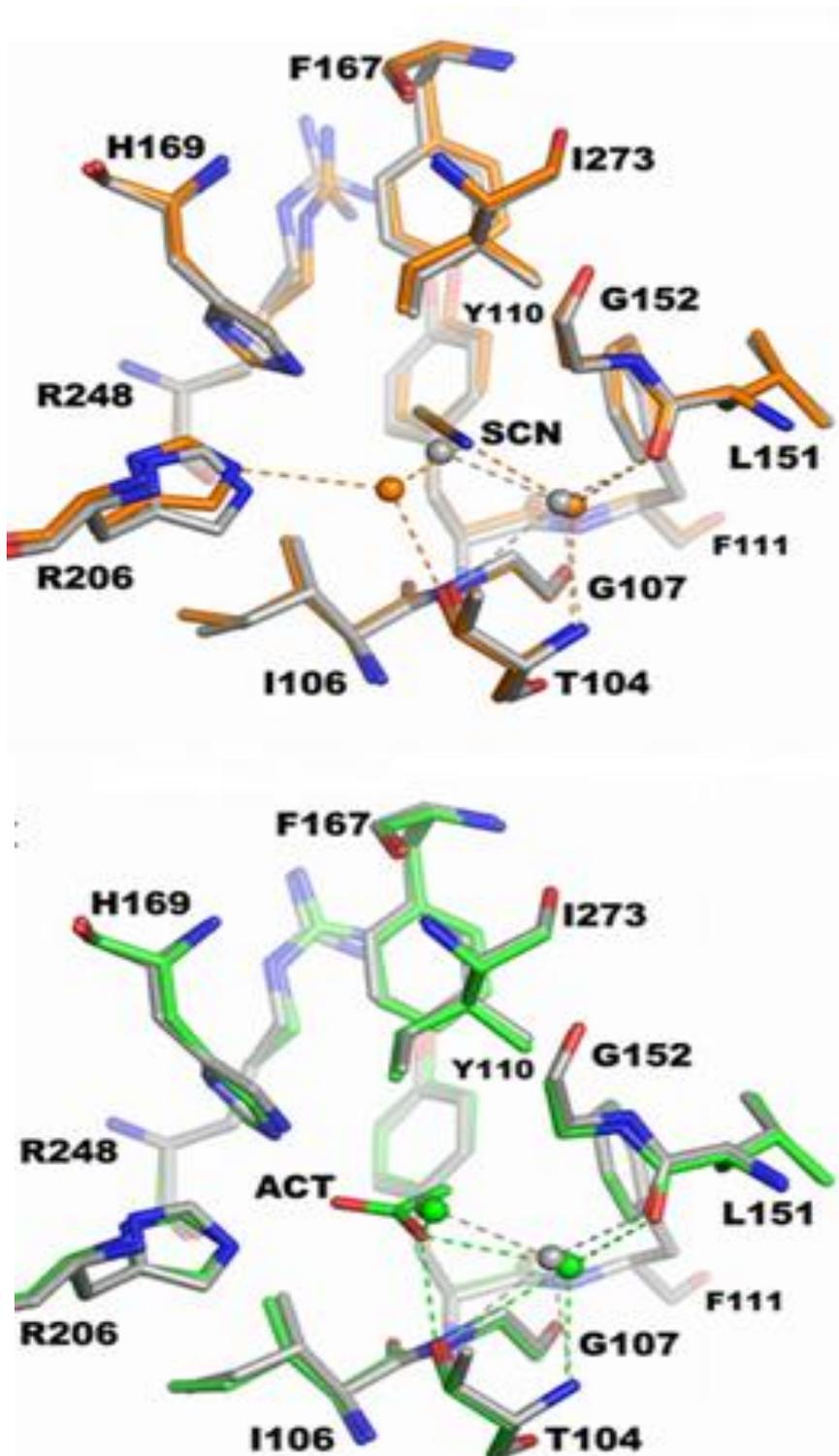
s

**Figure 3.3: A superposition of the crystal structures of homotetrameric *apo*-DntR and thiocyanate-bound DntR.** Top: Orthogonal views of a superposition of the full length tetramers. The monomers of *apo*-DntR are shown in different colors and thiocyanate-bound DntR in black. The r.m.s deviation in  $C_{\alpha}$  positions is 1.04 Å. Bottom: A superposition of the IBD tetrameric cores, the monomers of *apo*-DntR are shown in green and yellow, those of thiocyanate-bound DntR in black.

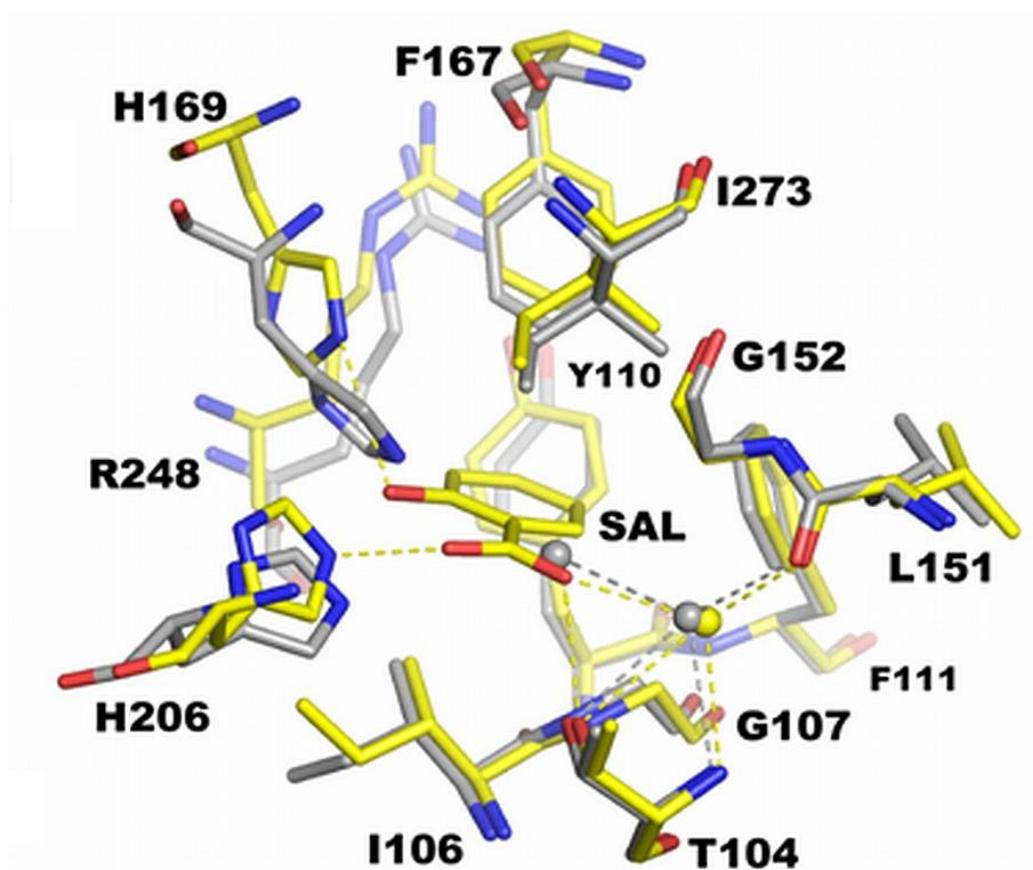
That the crystal structure of *apo*-DntR and acetate-/thiocyanate-bound DntR are so similar suggests that, in contrast to what had previously been proposed (Smirnova et al., 2004), the crystal structures of acetate and thiocyanate-bound DntR do not represent DntR homotetramers in an active conformation but rather in an inactive conformation. This idea is further validated by a comparison of the conformation of the IBCs observed in *apo*-, thiocyanate- and acetate-bound crystal structures (Figure 3.5). These are virtually identical. It could even appear that the acetate and thiocyanate ions are reinforcing the *apo* state of the IBC by forming direct or solvent mediated hydrogen bonds with amino acid residues H169 and H206 which are required for salicylate binding (Devesse et al., 2011). Figure 3.6 shows a comparison of the conformation of the IBCs observed in the crystal structures of *apo*-DntR and *holo*- $\Delta$ N90-DntR (Devesse et al., 2011). This makes it clear that with residues H169 and H206 locked in position by either acetate or thiocyanate DntR cannot bind salicylate as this requires an enlarged IBC formed by the displacement of the residues, particularly H169, that make up the IBC.



**Figure 3.4:** Electron density for the IBC in the crystal structure of *apo*-DntR.  $2mF_o - DF_c$  electron density, at the end of the refinement procedure, in the region of the IBC in the crystal structure of *apo*-DntR contoured at the  $0.9 \times r.m.s$  level.



**Figure 3.5:** A superposition of IBCs of the crystal structures of tetrameric *apo*-DntR and of acetate-/thiocyanate-bound DntR. Top: Superposition of the IBCs seen in the crystal structures of *apo*-DntR (grey carbon atoms) and thiocyanate-bound DntR (orange carbon atoms). Bottom: Superposition of the IBCs of *apo*-DntR (grey carbon atoms) and acetate-bound DntR (green carbon atoms).



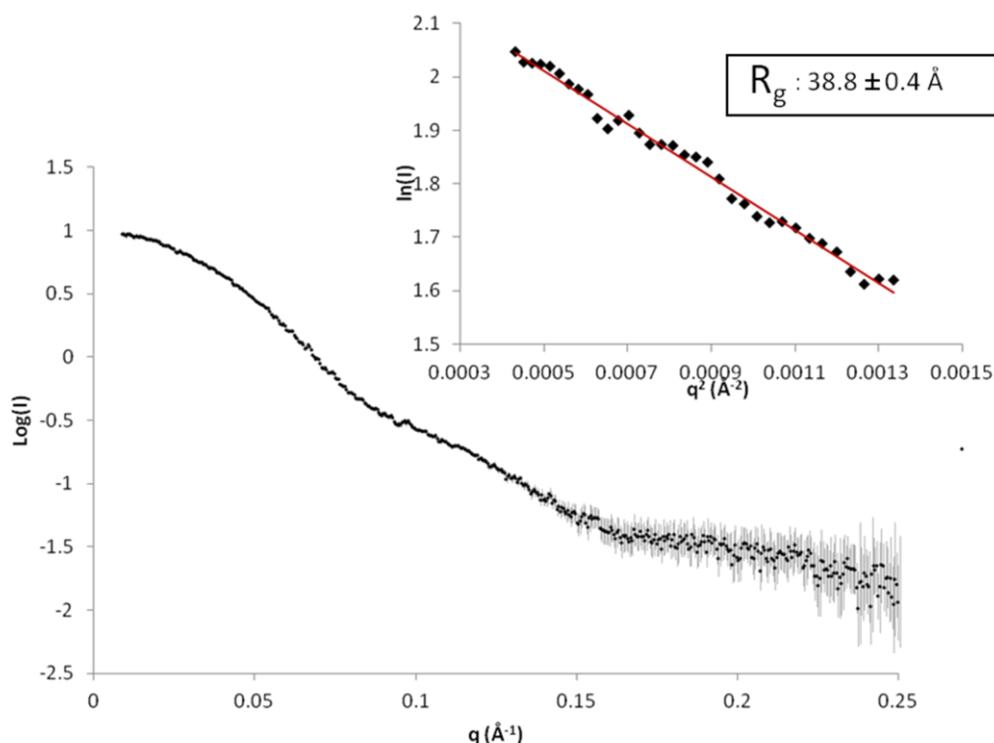
**Figure 3.6:** Comparison of the conformations of the IBCs of *apo*-DntR and *holo*- $\Delta$ N90-DntR. A superposition of the IBCs seen in the crystal structure of *apo*-DntR (grey carbon atoms), and salicylate-bound  $\Delta$ N90-DntR (yellow carbon atoms) (PDB 2Y7W). The Figure shows that a displacement of residues H169, H206 and R248 is required for salicylate binding.

### 3.1.2 Small Angle X-ray Scattering analysis *apo*-DntR

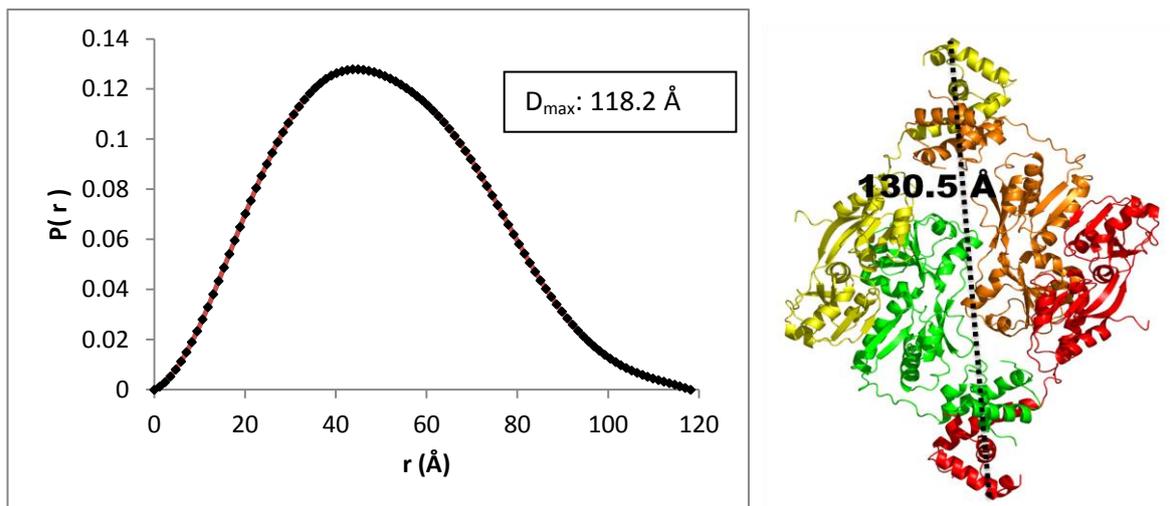
Although the crystal structure of *apo*-DntR appears to provide unequivocal evidence that *apo*-DntR homotetramers adopt a compact tetrameric structure, crystal packing effects might produce artefacts (Dickerson et al., 1994), that do not allow *apo*-DntR to adopt its in-solution conformation. To examine the behaviour of *apo*-DntR in solution SAXS experiments were performed. As outlined in Materials and Methods, DntR has a tendency to form higher order multimers. This complication made SAXS measurements at high concentrations of DntR unfeasible. Nevertheless, it was possible to obtain a satisfactory scattering curve containing no interparticle effects, combining measurements from solutions with *apo*-DntR concentrations of 3, 1 and 0.5 mg /ml (Figure 3.7). In the SAXS experiments a minimum of 17% (v/v) glycerol was added to the protein buffer to prevent inter-particle effects. Glycerol

reduces the contrast in electron density between the sample and solvent resulting in lower intensity which, in turn, reduces the apparent  $I(0)$  value. This makes molecular weights determined using  $I(0)$  unreliable (Putnam et al., 2007). This value is thus not used in any conclusions made concerning the solution structure of *apo*-DntR.

Analysis of the Guinier region of the scattering curve yielded a  $R_g$  value of 38.8 Å. The pair distribution function ( $P(r)$ ) derived from the scattering curve (Figure 3.8) indicates a maximum intra molecular distance of 118.2 Å, which is significantly shorter than the maximum distance (130 Å) observed in the crystal structure of full length *apo*-DntR homotetramers (Figure 3.8). CRY SOL fits (Figure 3.9) of the experimental scattering curve and theoretical scattering curves calculated from the crystal structure of *apo*-DntR and the open form DntR homology model based on the crystal structure of TsaR (Section 3.4.3) did not provide good fits and suggested that the solution structure of *apo*-DntR is significantly different to that observed in the crystal structures of both DntR and in the homology model of open-form DntR.



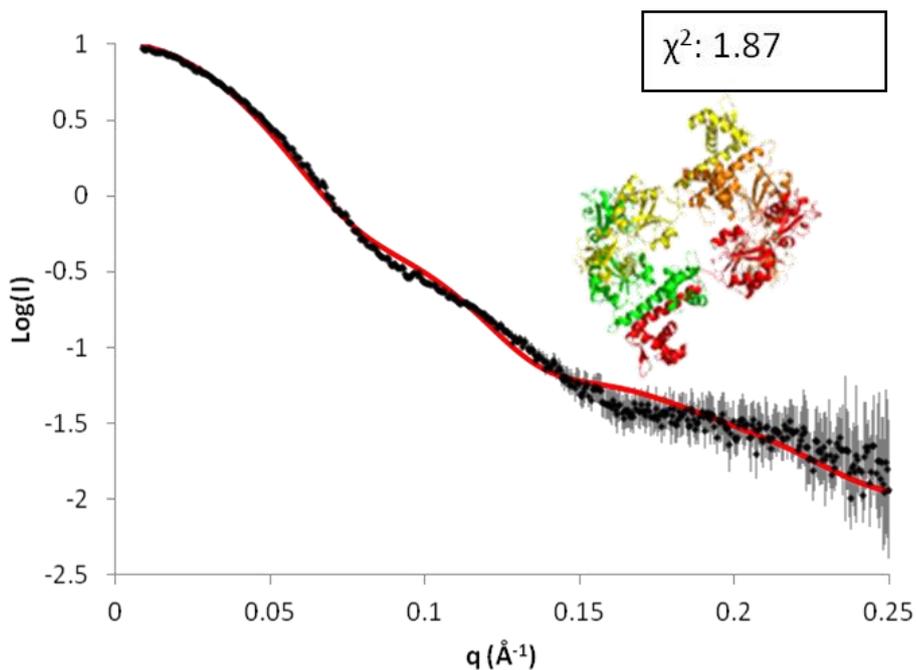
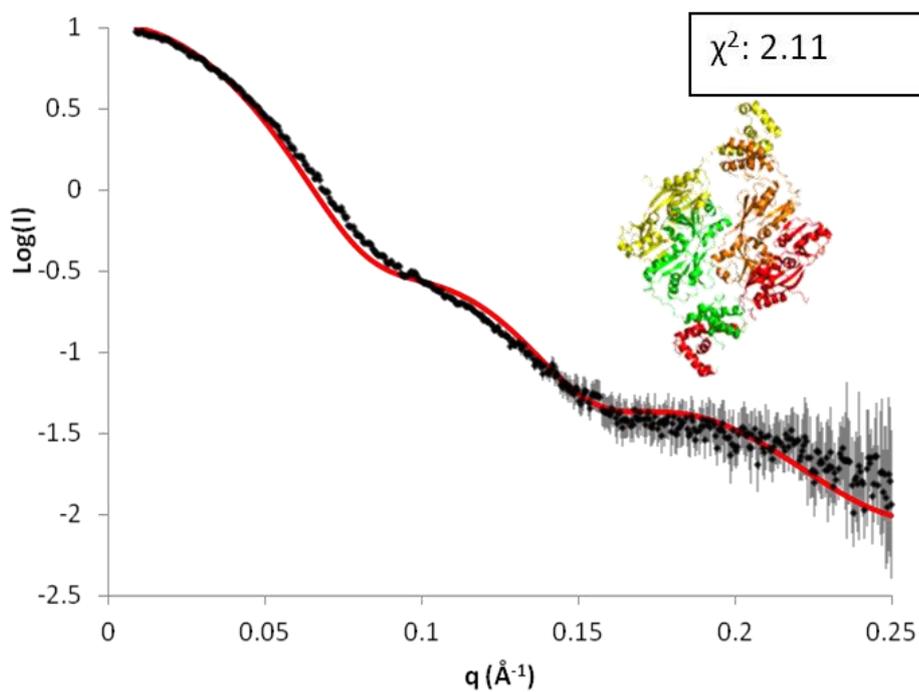
**Figure 3.7: The solution scattering curve and corresponding Guinier region for *apo*-DntR.** Solution scattering curve displaying  $\log(I)$  as a function of  $q$ . The curve displayed is the result of merging scattering curves measured at concentrations of 3, 1, and 0.5 mg/ml. Insert: Guinier region data points (black) and the resulting fit (red) derived from AUTOGNOM which yielded  $R_g = 38.8$  Å.



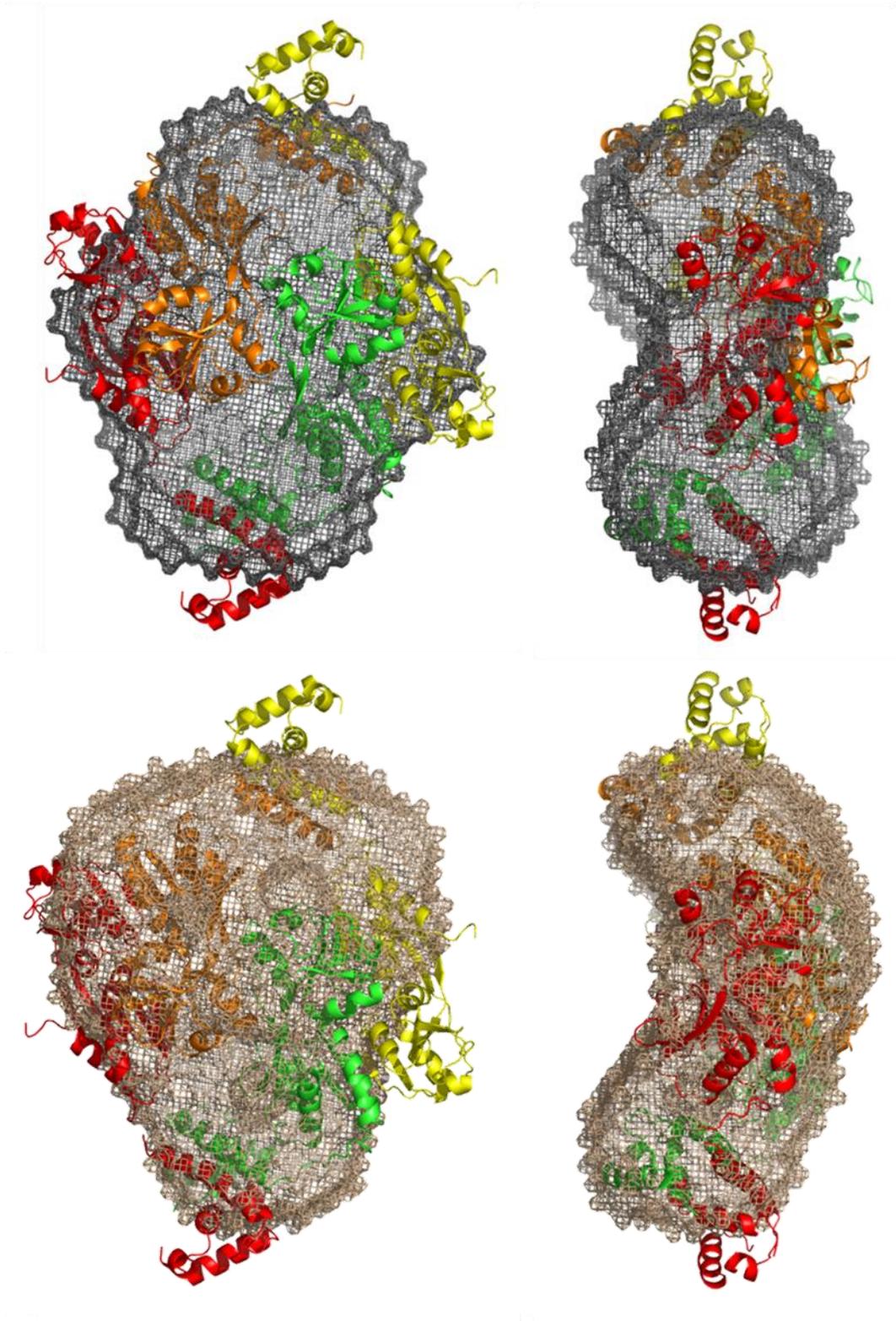
**Figure 3.8: The  $P(r)$  function of the solution structure of *apo*-DntR.** Left: SAXS-derived  $P(r)$  function obtained from the scattering curve shown in Figure 3.7 by indirect Fourier transformation using the program GNOM. Right: The crystal structure of the *apo*-DntR homotetramer with the longest intra molecular distance, as measured in PyMOL, displayed.

### 3.1.2.1 *ab initio* modelling

*ab initio* molecular envelopes of the solution structure of *apo*-DntR were created in DAMMIF (Franke and Svergun, 2009) based on the  $P(r)$  function shown in Figure 3.8. Models were created both without symmetry restraints and with P22 symmetry imposed. In both cases 20 independent models were obtained and averaged with DAMAVER. The results are shown in Figure 3.10. As expected from the  $P(r)$  function, the envelopes obtained indicate that the solution structure of *apo*-DntR is less elongated than that seen in the crystal structure described in Section 3.1.1.



**Figure 3.9: SAXS analysis of *apo*-DntR - CRY SOL fits of experimental and theoretical scattering curves.** To avoid any interference from interparticle effects the first few data points were removed. Top: CRY SOL fit of the experimental scattering curve obtained for *apo*-DntR and a theoretical curve based on the crystal structure of *apo*-DntR described in Section 3.1.1. Bottom: CRY SOL fit of the experimental scattering curve obtained for *apo*-DntR and a theoretical curve based on a homology model of open *apo*-DntR (see Section 2.4.3).



**Figure 3.10: SAXS-derived *ab initio* molecular envelopes for the solution structure of the *apo*-DntR homotetramer overlaid with the crystal structure of the *apo*-DntR homotetramer.** Models were based on 20 independent DAMMIF runs averaged and filtered with DAMAVER. The filtered models are shown. Enveloped were generated with imposed P22 symmetry (grey mesh, top) and without symmetry restrictions (brown mesh, bottom). The crystal structure of *apo*-DntR was manually superposed in PyMOL.

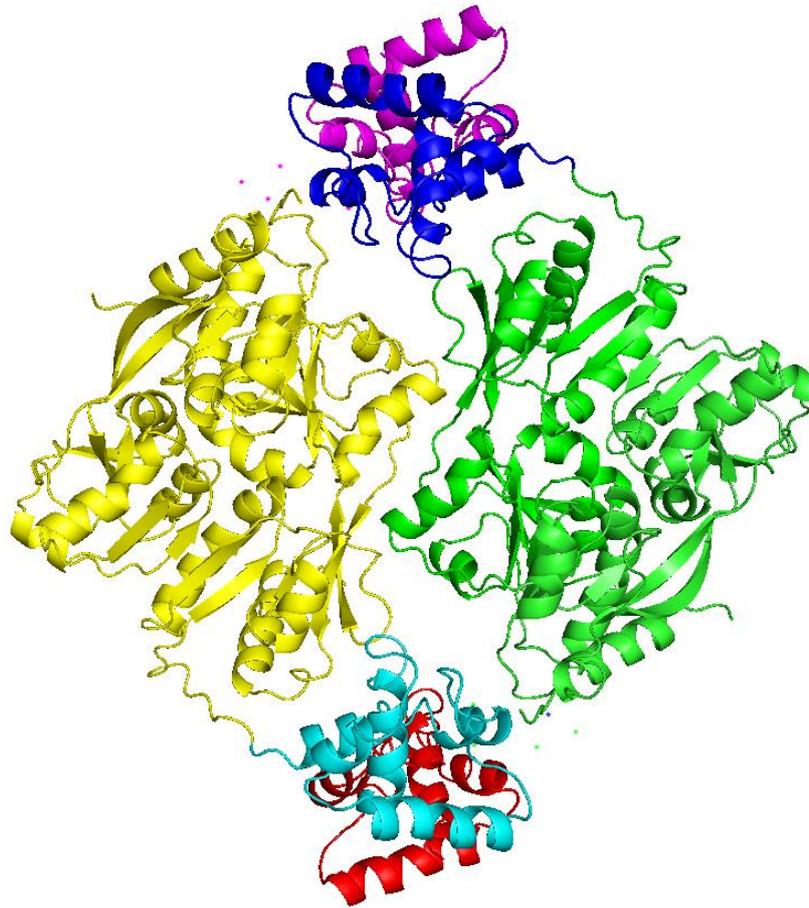
### 3.1.2.2 Rigid body modelling

In order to obtain a model of the solution structure of *apo*-DntR independent of bias introduced through the modelling of  $P(r)$  functions, rigid body refinement using SASREF was employed. The refinement protocol was based on the crystal structure of the *apo*-DntR described earlier (Section 3.1.1). Rigid bodies (Figure 3.11) were defined as described in Section 2.4.3.2. Each of the two IBD dimers (green and yellow in Figure 3.11) making up the compact DntR tetrameric core were treated as single rigid bodies; the 4 N-terminal wHTH domains were also each treated as individual rigid bodies. The division of the N and C terminal domains was set between residues T85 and T86 and a connectivity restraint of 3 Å was used to ensure that the two domains in each monomer remained connected. To ensure that the two wHTH domains making up a single dimeric DBD remained close together in space a connectivity restraint of 10 Å between residues A75 in the two monomers making up the dimer was employed.

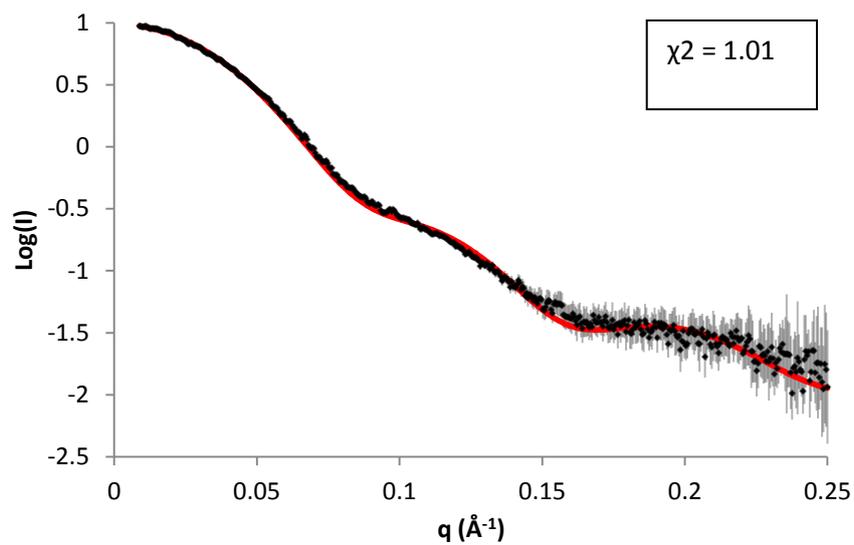
The result of the SASREF procedure is shown in Figure 3.12 and Table 3.2. As can be seen, the resulting model provides a much better fit to the *ab initio* envelope obtained (Figure 3.13) and, as expected from values of  $D_{\max}$  is less elongated in solution than *apo*-DntR homotetramer as seen in its crystal structure (Figure 3.14). In particular the wHTH dimers flanking the tetrameric core are packed closer to this than in the crystal structure. As previously demonstrated by the disordered electron density for these domains in crystal structures of homotetrameric DntR (Smirnova et al., 2004; this work) they are flexible and it therefore seems likely that their positions in crystal structures are a result of crystal packing trapping them in a state which is not their preferred in-solution position.

Model	$\chi^2$
SASREF generated solution structure	1.010
<i>apo</i> -DntR Crystal structure	2.114
Open-form homology model	1.870

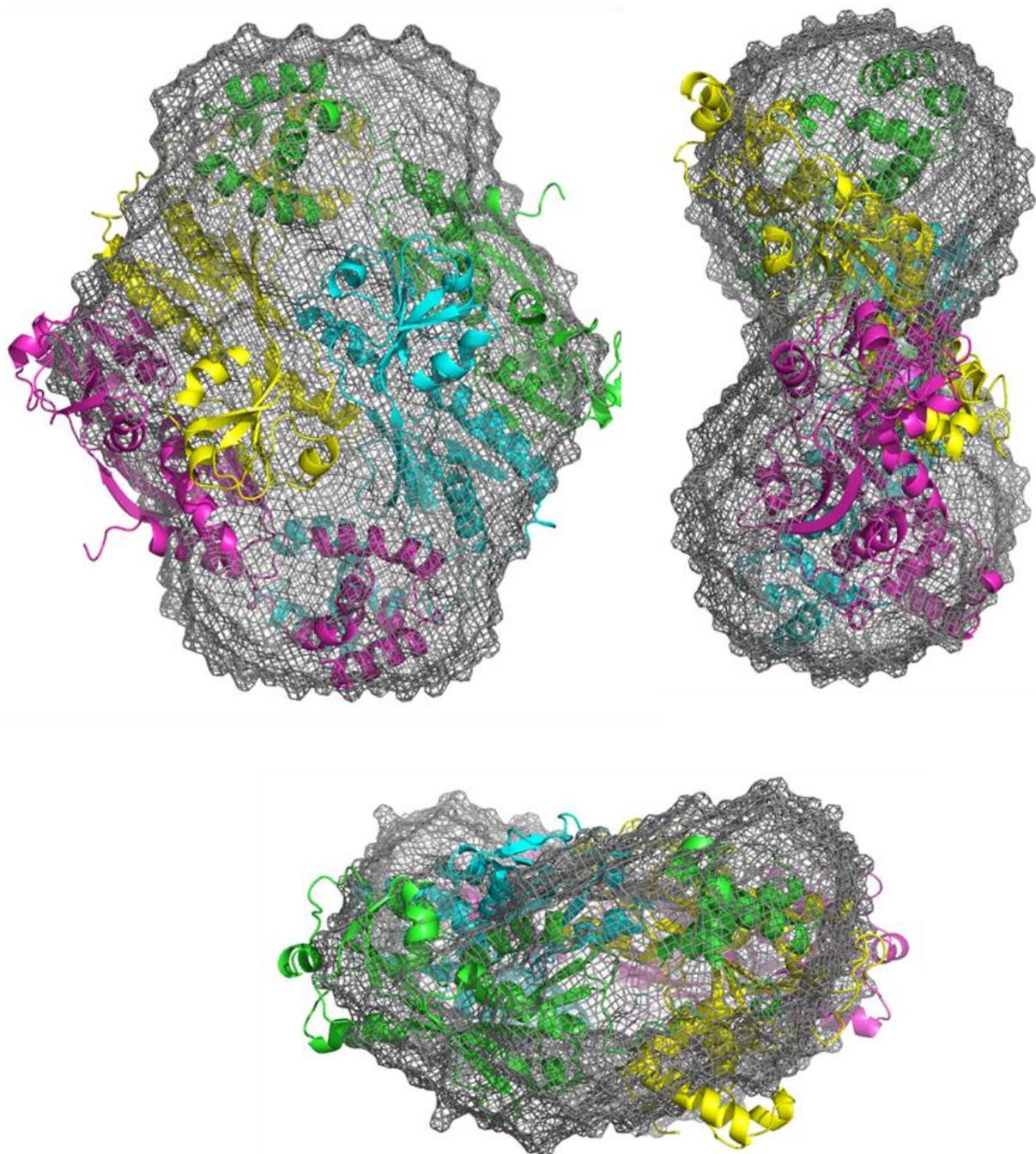
**Table 3.2:**  $\chi^2$  values of the CRY SOL fits shown in Figure 3.9 and 3.12.



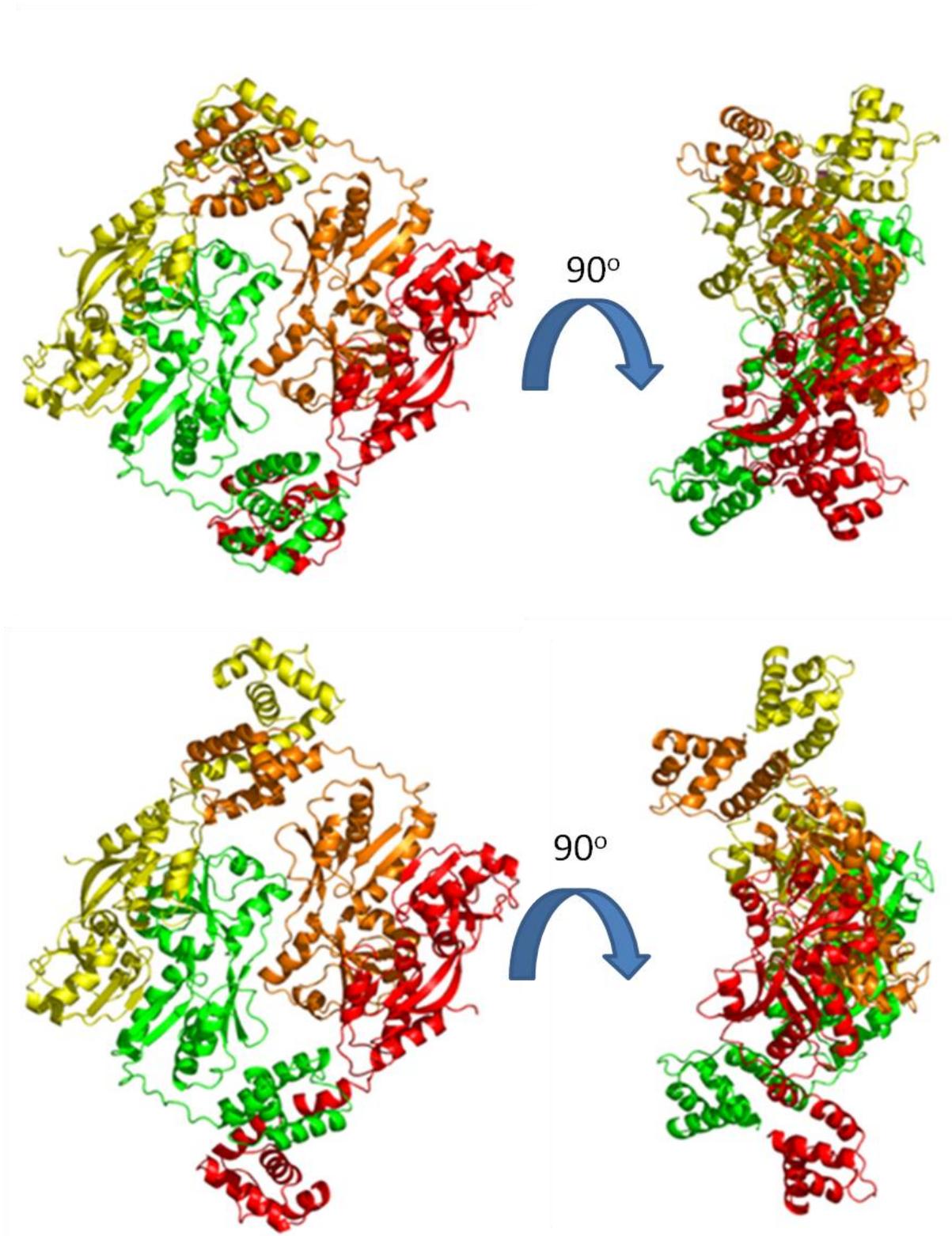
**Figure 3.11:** The rigid body composition of the crystal structure of DntR used in SASREF. The six rigid bodies used are shown in different colors.



**Figure 3.12:** Calculated scattering of the SASREF generated solution structure of *apo*-DntR vs. the experimental data. The theoretical scattering curve of the model generated (red) fitted against the experimental data (black).



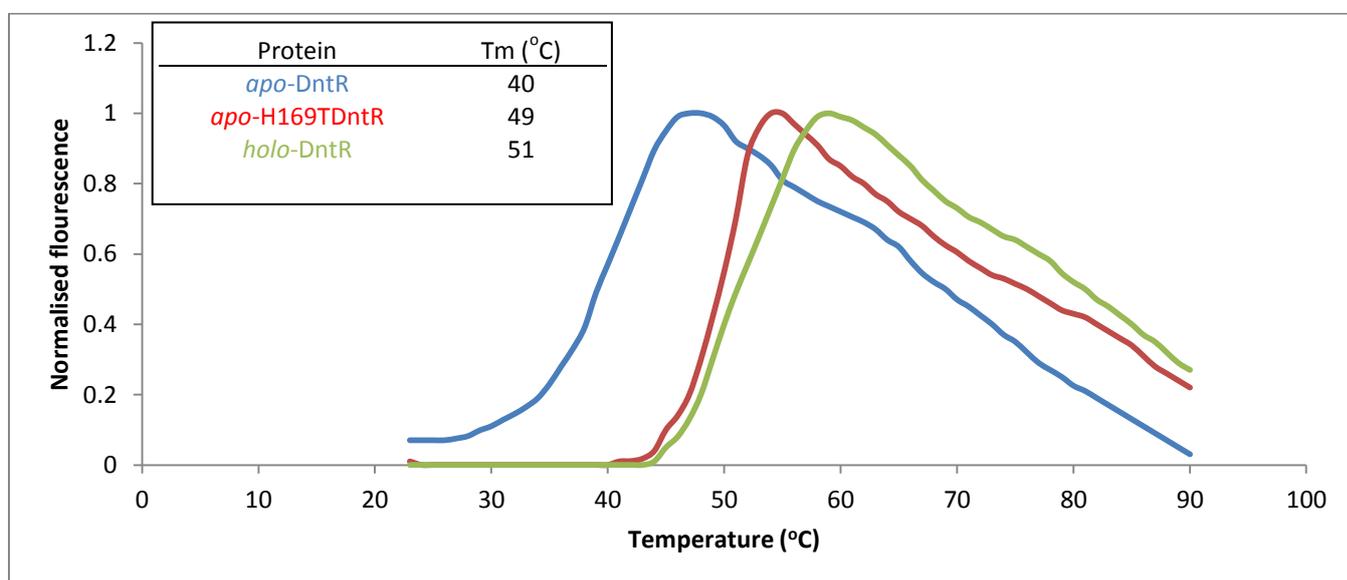
**Figure 3.13: SASREF-derived solution structure and *ab initio* envelopes for homotetrameric *apo*-DntR.** Three orthogonal views of the solution structure of homotetrameric *apo*-DntR as generated by SASREF overlaid with *ab initio* envelopes generated by DAMMIF and DAMAVER. The *ab initio* models are those created with imposed p22 symmetry.



**Figure 3.14: Orthogonal views of solution and crystal structures of *apo*-DntR.** Top: Orthogonal views of the solution structure of the *apo*-DntR homotetramers generated by SASREF. Bottom: Orthogonal views of the Crystal structure of full length *apo*-DntR. The major difference between the two is a clear change in the position of the wHTH dimers flanking the central IBD core.

## 3.2 H169TDntR

It has previously been shown that the H169T mutation causes DntR to be active even in the absence of an inducer molecule (Lónneborg, 2011, PhD thesis, University of Stockholm, Sweden). In this work Thermofluor assays carried out by the EMBL HTX platform (<https://embl.fr/htxlab>) (Figure 3.15) show that the H169T mutation causes a significant increase in the melting temperature of H169TDntR compared to that of *apo*-DntR, consistent with a difference in conformation. It was therefore postulated that the H169T mutant of DntR might adopt an active conformation and it was therefore decided to study the structure of this mutant both in crystals and in solution.



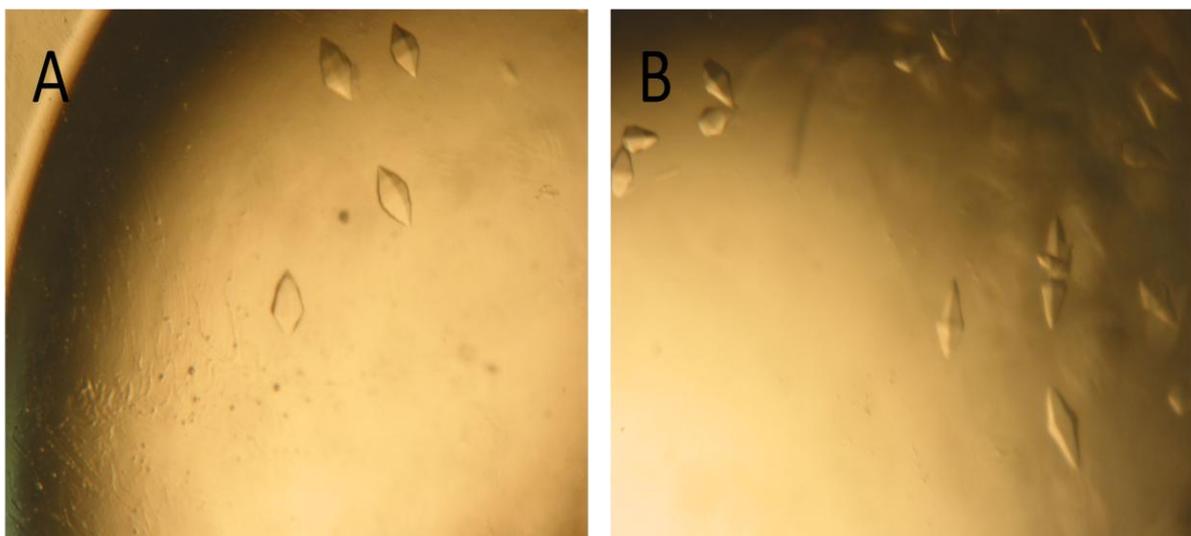
**Figure 3.15: Thermofluor assays of *apo*- and *holo*-DntR and *apo*-H169TDntR.** Introducing the H169T mutation caused a shift in melting temperature of DntR. The assays were conducted by the EMBL high throughput crystallisation laboratory (<https://embl.fr/htxlab/>). All protein concentrations were ~5 mg/ml. In the case of *holo*-DntR 100  $\mu$ M sodium salicylate was added to the sample.

### 3.2.1 Crystallization of H169TDntR

Crystals of the constitutively active H169T mutant of DntR (H169TDntR) could be produced in either of the conditions that produced *apo*-DntR and thiocyanate-bound DntR (Section 2.3.1).

However, crystals grown in the absence of thiocyanate (Figure 3.16B) diffracted only to  $d_{\min} \sim 12$  Å and were not pursued further (see Section 2.3.2). Those grown in the presence of

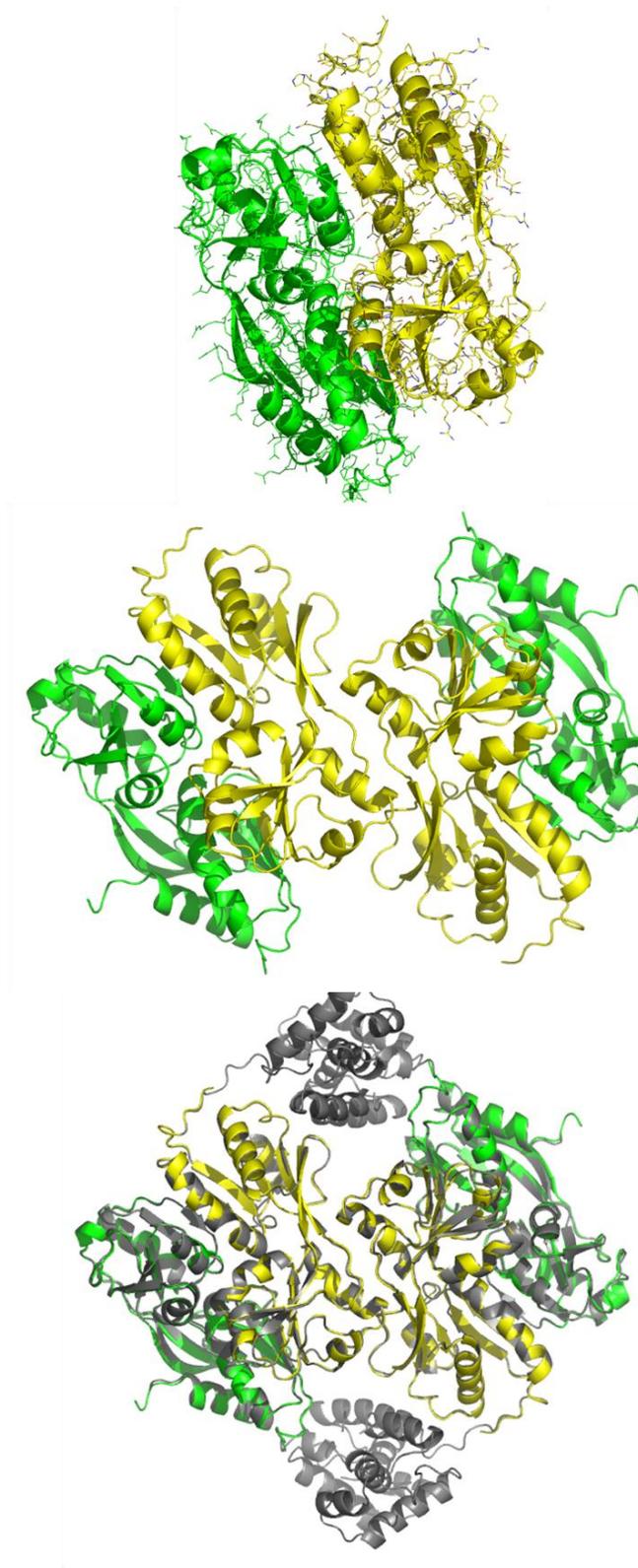
thiocyanate (Figure 3.16A) showed diffraction to  $d_{\min} \sim 3.3 \text{ \AA}$ . Full diffraction data collection, structure solution and refinement (Table 3.3) were therefore carried out using a single crystal of thiocyanate-bound H169TDntR. As can be seen from Table 3.3 the crystals of thiocyanate-bound H169TDntR obtained are isomorphous to those obtained for *apo*-DntR (this work), thiocyanate- or acetate-bound DntR (Smirnova et al., 2004) and other single point mutants of DntR (Lönneborg et al., 2007). The crystal structure of H169T is virtually identical to these. As for the crystal structure determined in (Lönneborg et al., 2007; Smirnova et al., 2004) the electron density for the wHTH dimers that flank the central core is very weak. These regions were thus not modelled and the crystal structure of H169TDntR thus consists of the central IBD tetrameric core which as for other crystal structure of full-length DntR homotetramers adopts a compact conformation (Figure 3.7).



**Figure 3.16: Crystals of H169TDntR.** (A): Crystals H169TDntR grown in the presence thiocyanate. (B): Crystals of H169TDntR grown in the absence of thiocyanate.

Dataset	H169T DntR
Beam-line	ID23-1 (ESRF)
Wavelength	0.976
Space group	P 6 <sub>3</sub> 22
Unit cell dimensions:	
a,b,c (Å)	107.47 107.47 299.77
α,β,γ (°)	90.00 90.00 120.00
Resolution range (Å)	47.14 - 3.30
Number of unique reflections	16123
Multiplicity	5.5 (5.6)
Completeness (%)	99.7 (100)
Rmerge	0.179 (0.951)
<I/σI>	10.9 (2.6)
Wilson B factor	88.799
Refined model composition	
Molecule A	S89-R302
Molecule B	R87-E300
Thiocyanate Ions	2
Model quality indicators	
R-factor (%)	18.82
Free R-factor (%)	24.03
RMS deviations:	
Bonds (Å)	0.007
Angles (°)	0.972
B factors	
<B> protein	76.20
<B> thiocyanate	86.29
Ramachandran analysis (%)	
Favored regions	97.18
Generously allowed regions	2.82
Disallowed regions	0.0

**Table 3.3: Data collection and refinement statistics for the crystal structure of H169TDntR.** Figures in parentheses are for the highest resolution range for the data set (3.48-3.3 Å).



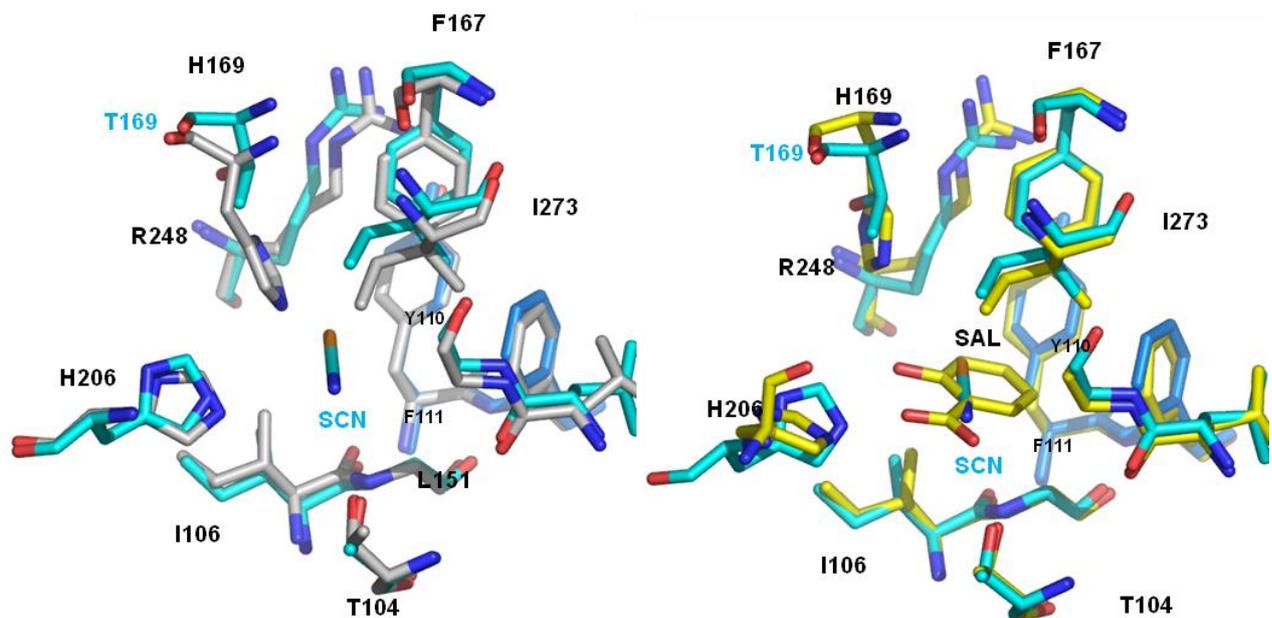
**Figure 3.17: The crystal structure of H169TDntR.** Top: The asymmetric unit observed in the crystal structure of H169TDntR. Middle: The structure of the H169TDntR IBD tetramer constructed based on symmetry related molecules such that they have coordinates  $(x, y, z)$  and  $(1-x, y, \frac{1}{2}z)$ . Bottom: Superposition of the crystal structure H169TDntR (yellow and green) and the solution structure of *apo*-DntR (black). The r.m.s. deviation in  $c_{\alpha}$  positions is 0.41 Å based on 431 residues aligned.

However, while the global structure of the IBD core of H169TDntR and *apo*-DntR are extremely similar, at a local level there is one very significant difference, specifically a superposition of the IBCs (Figure 3.18, left) seen in the crystal structures of *apo*-DntR and H169TDntR reveals that the H169T mutation results in a much more open binding site reminiscent of that observed in the IBC of salicylate-bound  $\Delta$ N90DntR (Figure 3.18, right). The crystal structure of H169TDntR thus allows for two possible interpretations:

- 1) The active form of DntR homotetramers maintains a closed tetrameric conformation but with more open IBCs.

However this does not explain the results of the thermofluor assays (Figure 3.15) which is consistent with different conformations for H169TDntR so another interpretation is:

- 2) Crystal packing forces H169TDntR to adopt an inactive homotetrameric state in the crystal form obtained.



**Figure 3.18:** Superposition of the IBCs seen in the crystal structures of *apo*-DntR, H169TDntR and *holo*- $\Delta$ N90DntR. Left: The IBC of H169TDntR (blue carbon atoms) superposed with IBC of *apo*-DntR (grey carbon atoms). Right: The IBC of H169TDntR (blue carbon atoms) superposed with IBC of salicylate-bound  $\Delta$ N90-DntR (yellow carbon atoms).

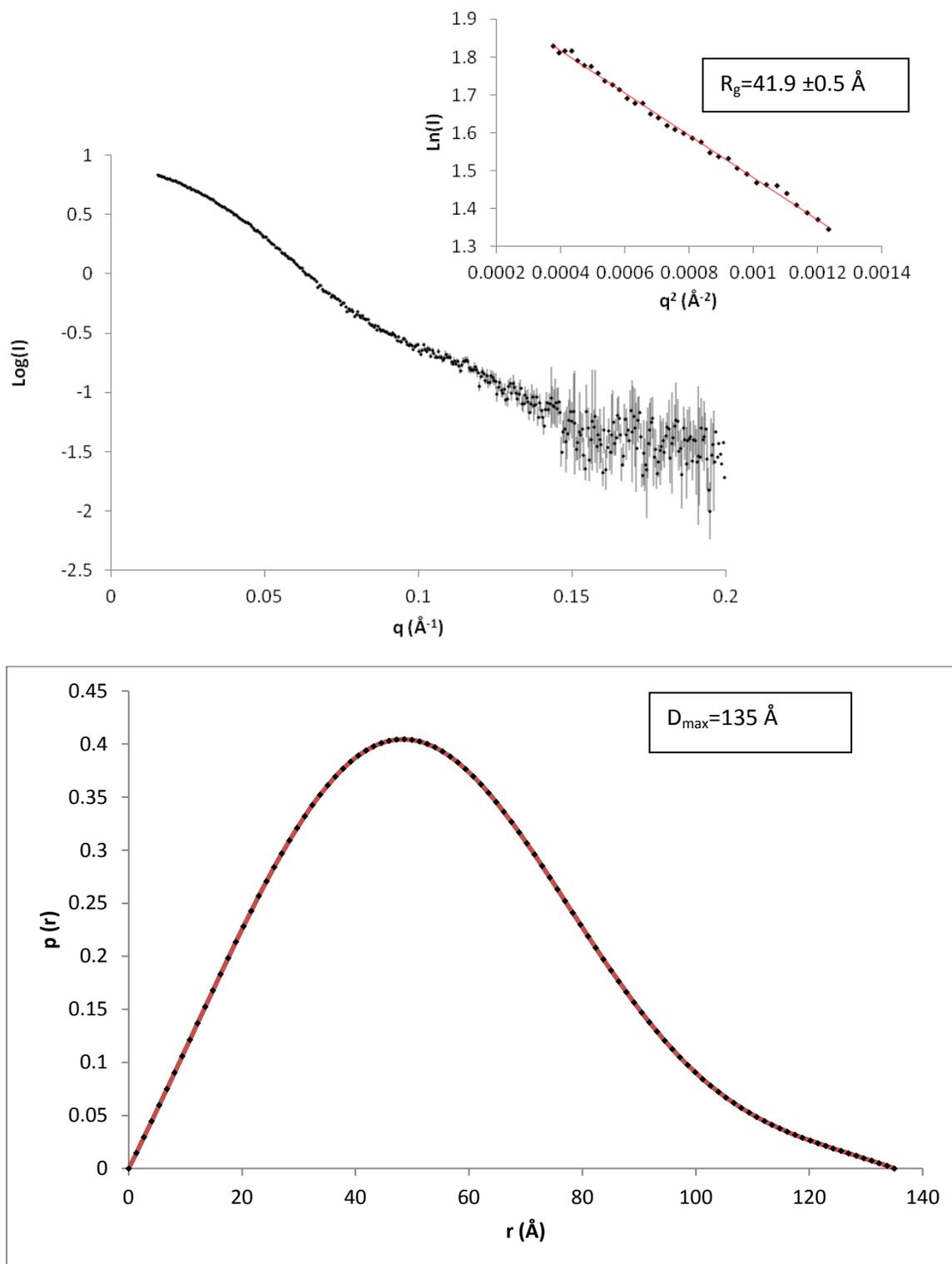
### 3.2.2 Small Angle X-ray Scattering analysis H169TDntR

To examine the structure of *apo*-H169TDntR in solution, SAXS experiments were performed (Section 2.4.1.3). H169TDntR showed a higher propensity to aggregate than did DntR and to avoid any interparticle interactions the concentration of solutions of H169T used in SAXS experiments was kept at 0.7 mg/ml. In order to achieve higher resolution data than otherwise would have been the case, 10 x 10 frames, each of for 2 seconds exposure time, were averaged to produce the experimental scattering curve (Figure 3.19, top).

Analysis of the Guinier region of the scattering curve yielded a  $R_g$  value of 42 Å (Table 3.4) which is significantly larger than that seen in the solution structure (38.8 Å) of the *apo*-DntR tetramer. The pair distribution ( $P(r)$ ) function derived from the scattering curve (Figure 3.19, bottom) indicates a  $D_{max}$  of 135 Å (Table 3.4), which is significantly longer than the maximum distance (118 Å) observed in the solution structure of full length *apo*-DntR (Figure 3.8) but is however a close fit to what was observed in the crystal structure of *apo*-DntR (130 Å).

Invariant Parameter	Solution structure	Solution structure	Crystal structure
	<i>apo</i> -DntR	<i>apo</i> -H169TDntR	<i>apo</i> -DntR
$D_{max}$ (Å)	118	135	130
$R_g$ (Å)	38.8 +/- 0.6	41.9 +/- 0.5	

**Table 3.4: Comparison of invariant parameters obtained from solution scattering curves for *apo*-DntR and *apo*-H169TDntR.**



**Figure 3.19: SAXS scattering curve and corresponding Guinier region and  $P(r)$  function for *apo*-H169TDntR.** Top: Scattering curve displaying  $\text{log}(I)$  as a function of  $q$ . Insert: Guinier region data points (black) and the resulting fit derived from AUTOGNOM (red) which yielded  $R_g = 41.9 \text{ \AA}$ . Bottom: SAXS-derived Pair Distribution Function of full length H169TDntR yielding  $D_{\text{max}} = 135$ . The function was generated by indirect Fourier transformation using the program GNOM.

### 3.2.2.1 Molecular envelopes of the solution structures of apo-H169TDntR

*ab initio* envelopes of the solution structure of apo-H169TDntR were created in DAMMIF based on the P(r) function shown in Figure 3.19. Twenty independent models were obtained, applying p22 symmetry restrictions, and averaged with DAMAVER. The result is shown in Figure 3.21. The envelope obtained suggests that in solution H169TDntR adopts much more elongated conformation than that seen for the solution structure of apo-DntR. Moreover, the envelope obtained is more consistent with the open-form homotetramer structure of the TsaR derived homology model (Figure 3.20, bottom). This observation is reinforced by the results of CRY SOL fits of the experimental scattering curves vs. the theoretical scattering curves based on other structural models.

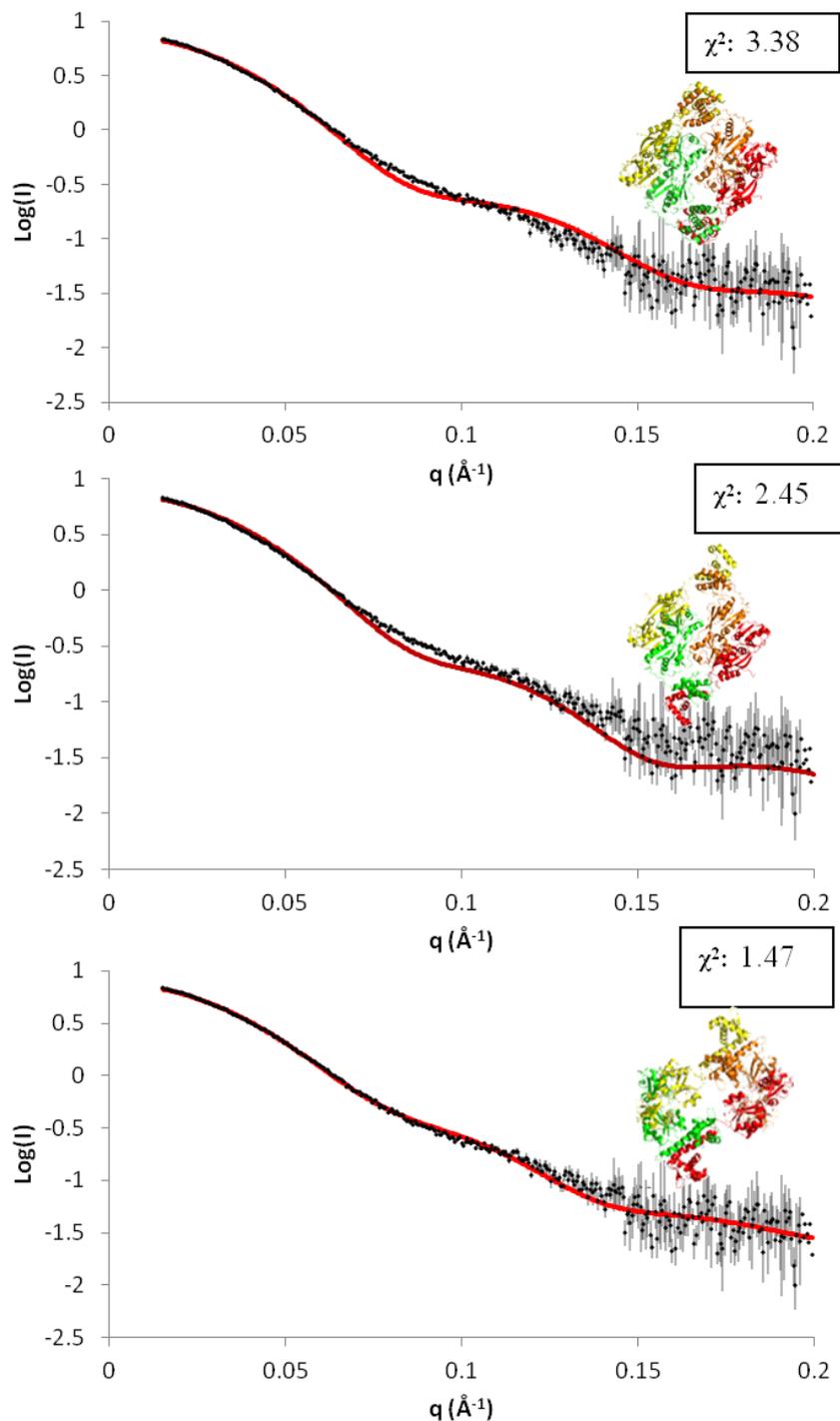
Model	$\chi^2$
Apo-Solution Structure	3.38
Apo-Crystal Structure	2.45
TsaR Homology Model	1.47

**Table 3.5:**  $\chi^2$  values obtained for the CRY SOL fits shown in Figure 3.20.

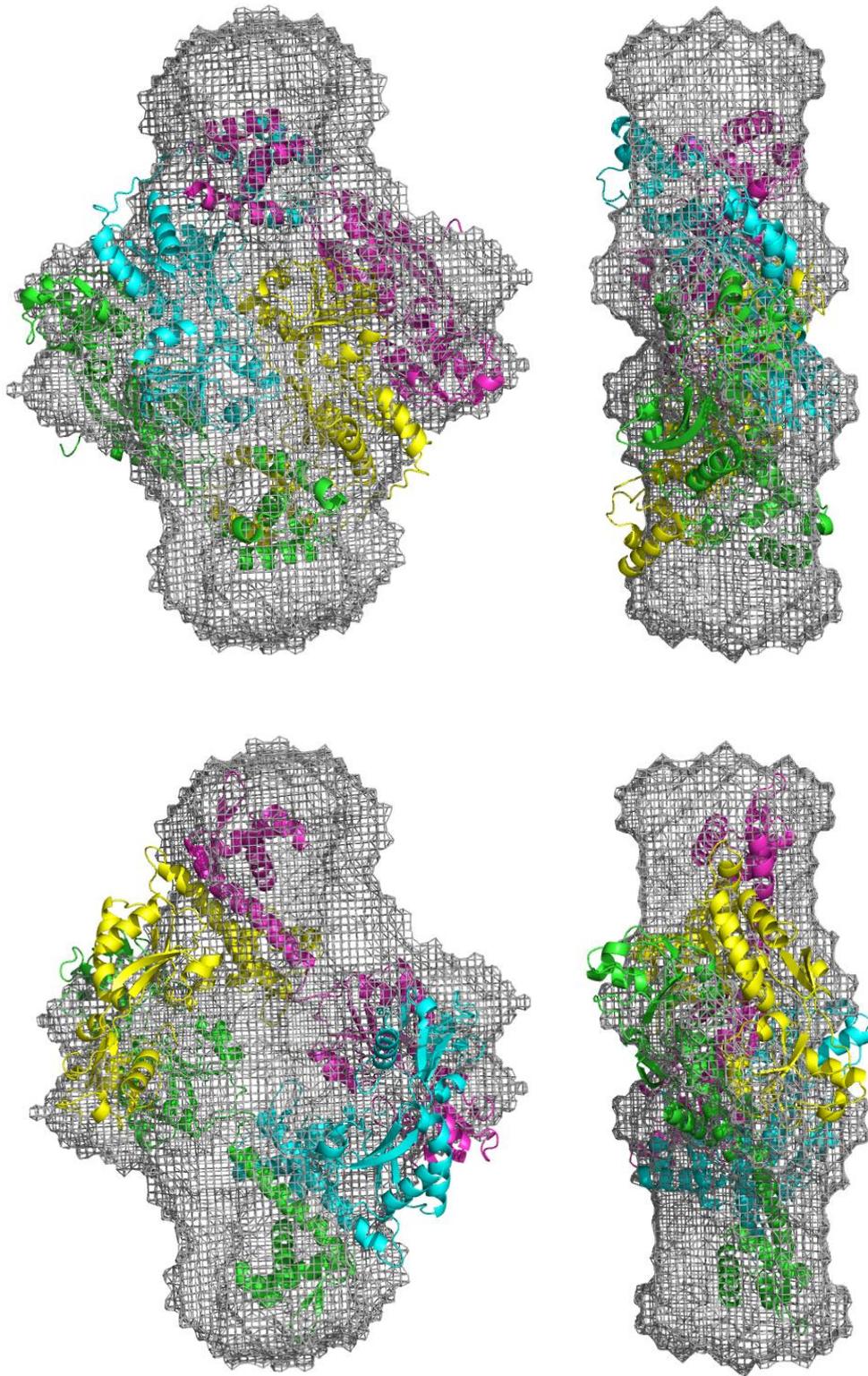
As can be seen in both Table 3.5 the calculated scattering curve based on the solution structure apo-DntR as described previously (Section 3.1.2) is a very poor fit to the experimentally obtained scattering curve for apo-H169TDntR. Moreover, while a calculated scattering curve based on the crystal structure of apo-DntR tetramer (Section 3.1.2.1) produces an improved, fit it is clear that a calculated scattering curve based on the open form DntR homology model based on the structure of TsaR model provides a better fit than the theoretical curves of both other structures.

Rigid body modelling was attempted but the scattering data proved insufficient as multiple runs did not converge towards a single model and modelling was therefore abandoned.

Nevertheless both the *ab initio* envelopes obtained and CRY SOL the fitting of an experimental and calculated scattering curve suggest that in solution H169TDntR tetramers adopt a open-form TsaR-like conformation.



**Figure 3.20: SAXS analysis of *apo*-H169TDntR - CRY SOL fits against the experimental scattering data.** To avoid any interference from interparticle effects the first few data points were omitted. Top: Theoretical scattering curve of Solution structure of *apo*-DntR as generated by SASREF (red) vs. the experimental scattering curve for *apo*-H169TDntR (black). Middle: Theoretical scattering curve of the crystal structure of full length *apo*-DntR (red) vs. the experimental scattering curve for *apo*-H169TDntR (black). Bottom: Theoretical scattering curve of the open-form DntR homology model (red) vs. the experimental scattering curve for *apo*-H169TDntR (black). The resulting  $\chi^2$  values of the fits are shown as insert is and in Table 3.6.



**Figure 3.21:** *ab initio* envelope of the solution structure of H169TDntR. Top: *ab initio* envelope manually overlaid with the previously determined (Section 3.2.2.2) solution structure of apo-DntR. Bottom *ab initio* envelope overlaid with open-form DntR homology model based on the crystal structure of TsaR.

### 3.2.2.2 Comparison of the results of SAXS experiments concerning *apo*-DntR and *apo*-H169TDntR

The results presented above highlight a case in which different proteins have the same crystal structure but have markedly different structures in solution. The SAXS studies described above clearly also show that while *apo*-DntR maintains a closed tetrameric conformation in solution, H169TDntR does not. For *apo*-H169T DntR SAXS derived parameters ( $R_g$ ,  $D_{max}$ ; Table 3.5) indicate a structure slightly more elongated than seen in the crystal structure. This evidence is further supported by *ab initio* models and CRY SOL fits of theoretical scattering curves to the experimental data which suggest that in solution the H169TDntR homotetramer has a structure close to that of the open form tetramer seen in the crystal structure of TsaR. That H169TDntR shows activity even in the absence of its inducer molecule salicylate suggests that it adopts an activated conformation. This suggests that the open conformation observed in TsaR is indeed the activated state of an LTTR protein. The results of the SAXS studies of *apo*-DntR and *apo*-H169TDntR thus appear to confirm the hypothesis (Devesse et al., 2011; Monferrer et al., 2010) that upon activation LTTR homotetramers change conformation from a compact homotetrameric state to a more open state.

### 3.3 Conclusions on inducer-free DntR

The solution structures presented herein show that DntR in its *apo* state does indeed present a compact conformation. The crystal structure of *apo*-DntR shows that the conformation of the IBC in the *apo* state prevents inducer molecules from binding. The current hypothesis (Devesse et al., 2011; Monferrer et al., 2010) of LTTR activation is that inducer binding results in a change of conformation of the IBC which translates into a movement of the hinge region in the IBDs causing a closure of the IBDs and a transition into a more extended homotetramer. This model seems to be confirmed by the crystallographic studies of H169TDntR presented here. X-ray crystallography shows that the IBCs in H169TDntR, which can activate transcription in the absence of an inducer molecule, adopt a conformation more similar to those seen in the IBCs of salicylate-bound DntR than the IBCs of *apo*-DntR. However, SAXS studies show that in solution H169TDntR homotetramers adopt an open and elongated conformation. The hypothesis that the solution structure of

H169TDntR is really that of an activated DntR homotetramer needs, however, to be confirmed by in solution studies of DntR in presence of its inducer molecule salicylate.

### 3.4 Small Angle X-ray Scattering analysis of DntR in complex with salicylate

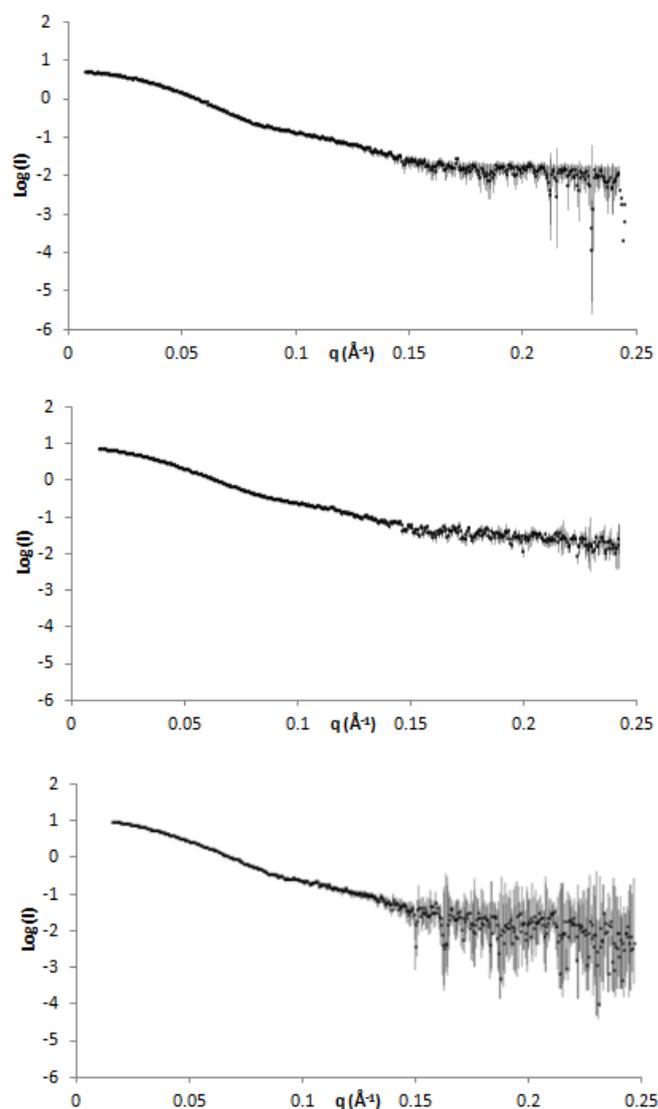
The SAXS results obtained from solutions of H169TDntR coupled with evidence that this mutant is constitutively active and thermofluor studies indicating that the mutant adopts a different conformation than *apo*-DntR (Figure 3.15) suggests that the active form of DntR is similar to the open form tetramers seen in the crystal structures of TsaR (Monferrer et al., 2010). To confirm this suggestion, SAXS scattering curves from solutions of DntR pre-incubated sodium salicylate in concentrations of 20, 100 and 5000  $\mu$ M were measured as outlined in Section 2.4. Measurements of the scattering curves from solutions of *apo*-DntR were also repeated and showed results similar to those previously obtained Section 3.1.2. The resulting scattering curves observed from solutions of DntR pre-incubated with sodium salicylate are shown in Figure 3.22. Invariant parameters ( $R_g$ ,  $I(0)$ ,  $D_{max}$ ) were obtained using  $AUTOR_G$  and  $AUTOGNOM$ . The results are displayed in Figure 3.22, 3.23 and 3.24 and Table 3.6. Although the  $I(0)$  values cannot be correlated with an absolute molecular mass due to high glycerol concentration (see Section 3.1.2) it can still be used arbitrarily to compare the different samples.

The addition of 20  $\mu$ M sodium salicylate shows no significant changes in invariant parameters compared to the values obtained from solutions of *apo*-DntR. Pre-incubation with 100  $\mu$ M sodium salicylate yields increased values of  $R_g$  and  $D_{max}$  compared to those of *apo*-DntR. However, the values of  $I(0)$  and Porod volume do not increase significantly suggesting that changes in this scattering curve compared to that obtained for *apo*-DntR can be attributed to structural differences.

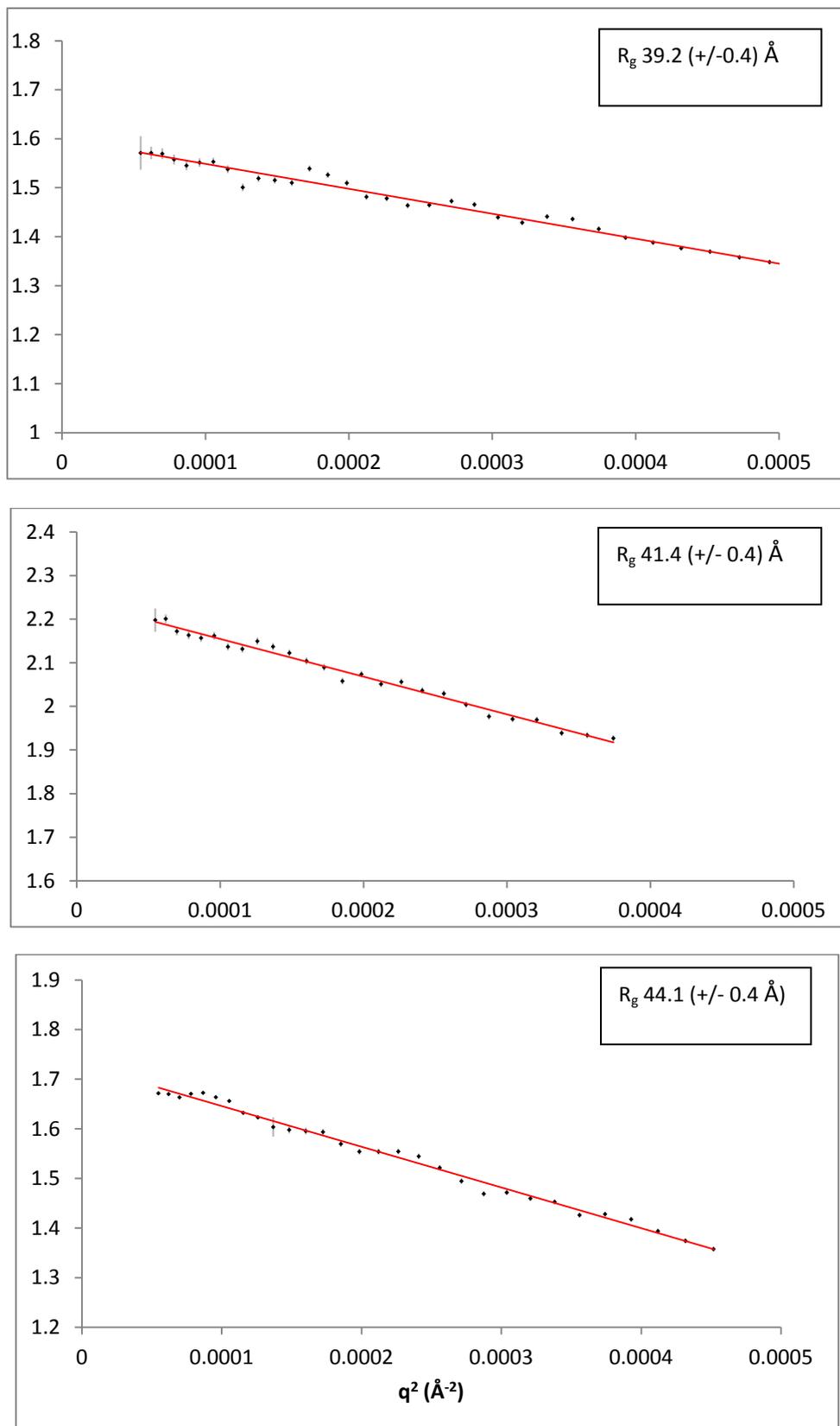
[salicylate] $\mu$ M	$R_g$ ( $\text{\AA}$ )	$I(0)$	$D_{max}$ ( $\text{\AA}$ )	Porod volume $\text{nm}^3$
0	38.4 (+/-0.6)	4.97 (+/- 0.04)	118	262.85
20	39.2 (+/-0.4)	5.18 (+/-0.02)	120	262.50
100	41.4 (+/- 0.4)	5.22 (+/-0.03)	142	262.19
5000	44.1 (+/- 0.4)	5.88 (+/-0.04)	147	296.89

**Table 3.6: Invariant parameters obtained from SAXS scattering curves from solutions of *apo*-DntR and of DntR pre-incubated with different concentrations of sodium salicylate.** Parameters obtained as a function of salicylate.

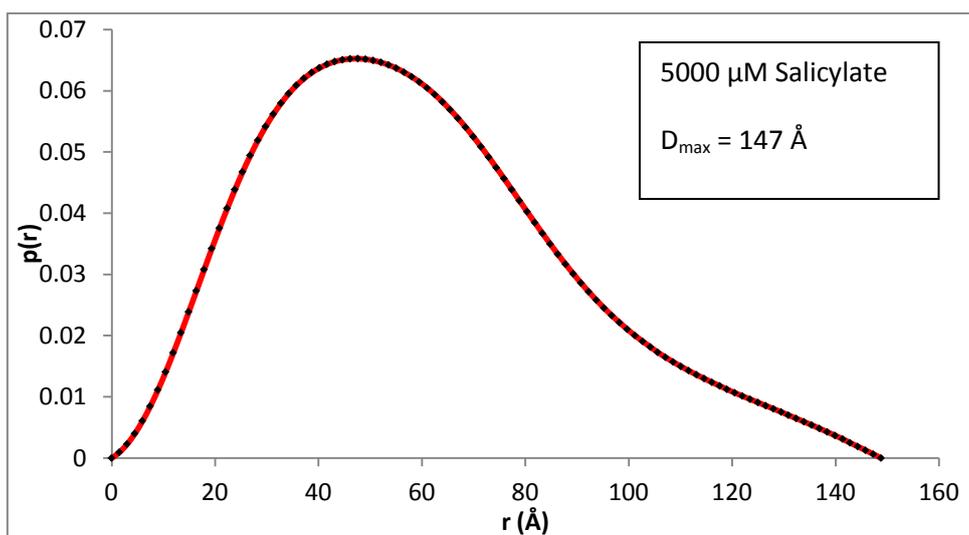
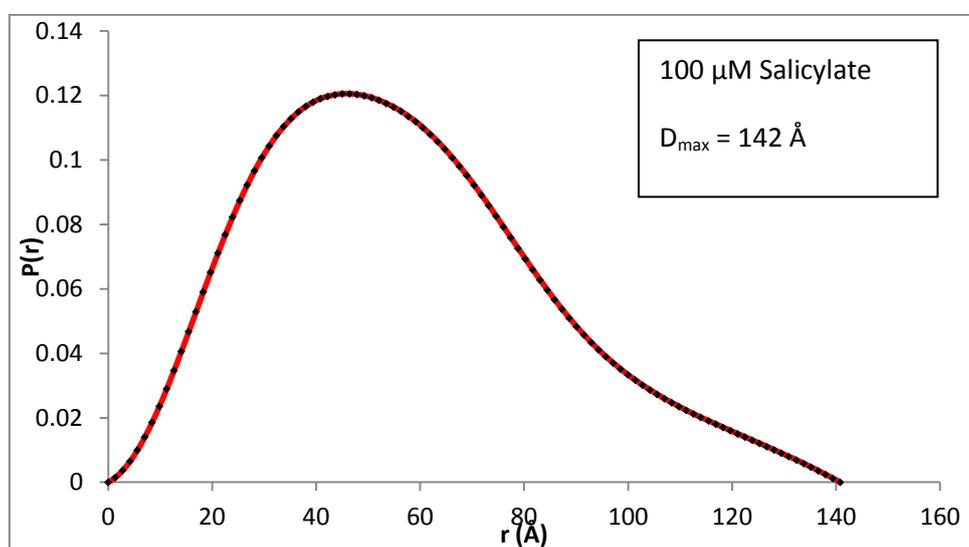
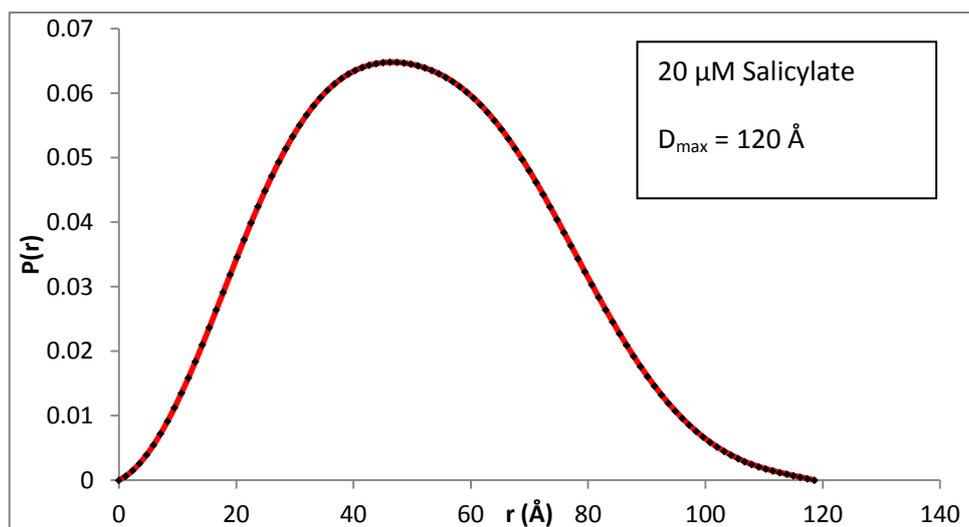
As pre-incubation with sodium salicylate should not significantly increase the molecular weight of the DntR sample under study (salicylate mw 138 Da), any large increase in  $I(0)$  and in Porod Volume (Table 3.6) is problematic as this may indicate interparticle interactions (Putnam et al., 2007). The large change in the Porod Volume and  $I(0)$  value upon addition of 5000  $\mu\text{M}$  sodium salicylate indicates the introduction of sufficient inter-particle effects to make this scattering curve unreliable for further analysis. Determination of the solution structures of salicylate-bound DntR was therefore based on SAXS scattering curves measured from a solution of DntR pre-incubated with 100  $\mu\text{M}$  sodium salicylate.



**Figure 3.22: Solution scattering curves obtained from solutions of DntR pre-incubated with various concentrations of sodium salicylate.** Scattering curve displaying  $\log(I)$  as a function of  $q$  for DntR pre-incubated with 20  $\mu\text{M}$  (top), 100  $\mu\text{M}$  (middle), and 5000  $\mu\text{M}$  (bottom) sodium salicylate. The curves are displayed within the data range  $q = 0\text{-}0.25 \text{ \AA}^{-1}$ .



**Figure 3.23: The Guinier regions of scattering curves obtained from solutions of DntR pre-incubated with salicylate.** Guinier regions data points (black) and the resulting fit (red) derived from AUTOGNOM for solutions of DntR pre-incubated with 20  $\mu\text{M}$  (Top), 100  $\mu\text{M}$  (Middle) and 5000  $\mu\text{M}$  (Bottom) sodium salicylate. The  $R_g$  values are shown in inserts.



**Figure 3.24:  $P(r)$  functions of the scattering curves shown in Figure 3.22.**  $P(r)$  functions of the scattering curves obtained in the presence of 20  $\mu\text{M}$  (Top), 100  $\mu\text{M}$  (Middle) and 5000  $\mu\text{M}$  (Bottom) sodium salicylate.

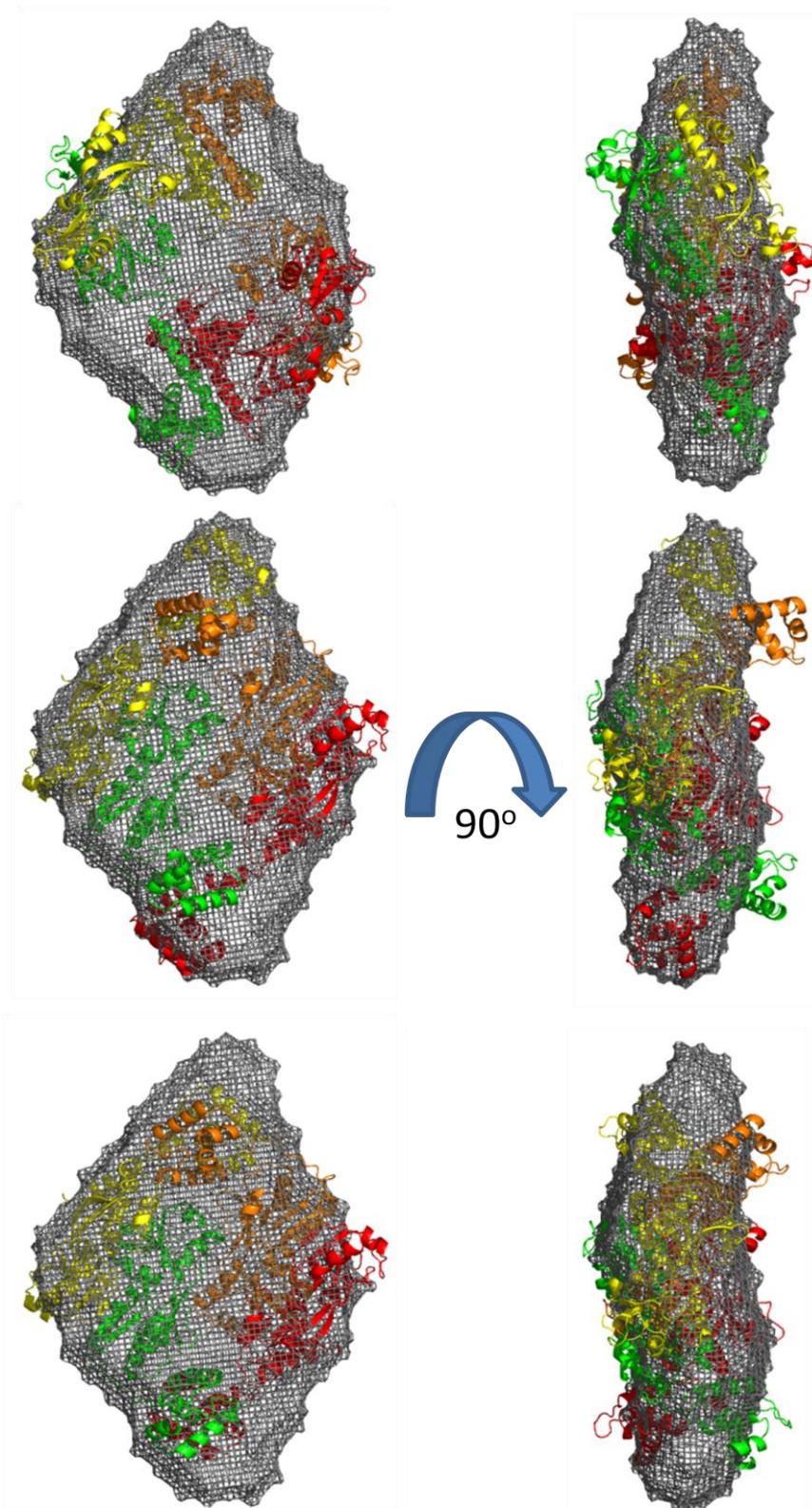
### 3.4.1 Solution structure of DntR pre-incubated with 100 $\mu$ M salicylate

The solution scattering curve measured from DntR pre-incubated with 100  $\mu$ M sodium salicylate was used to produce *ab initio* molecular envelopes (Figure 3.25) of the solution structure of salicylate-bound, and presumably activated, DntR. Here, to reduce bias, modelling was conducted without symmetry restrictions. The envelope obtained is clearly more similar in nature to the envelope obtained (Figure 3.21) for the H169TDntR homotetramer than that for the *apo*-DntR homotetramer (Figure 3.10). It is also evident that the crystal structure of the *apo*-DntR (Section 3.1) and the open-form TsaR-like homology model of a DntR homotetramer, described earlier (Section 3.4.3), fit the envelope much better than the solution structure of *apo*-DntR obtained in section 3.1.2.

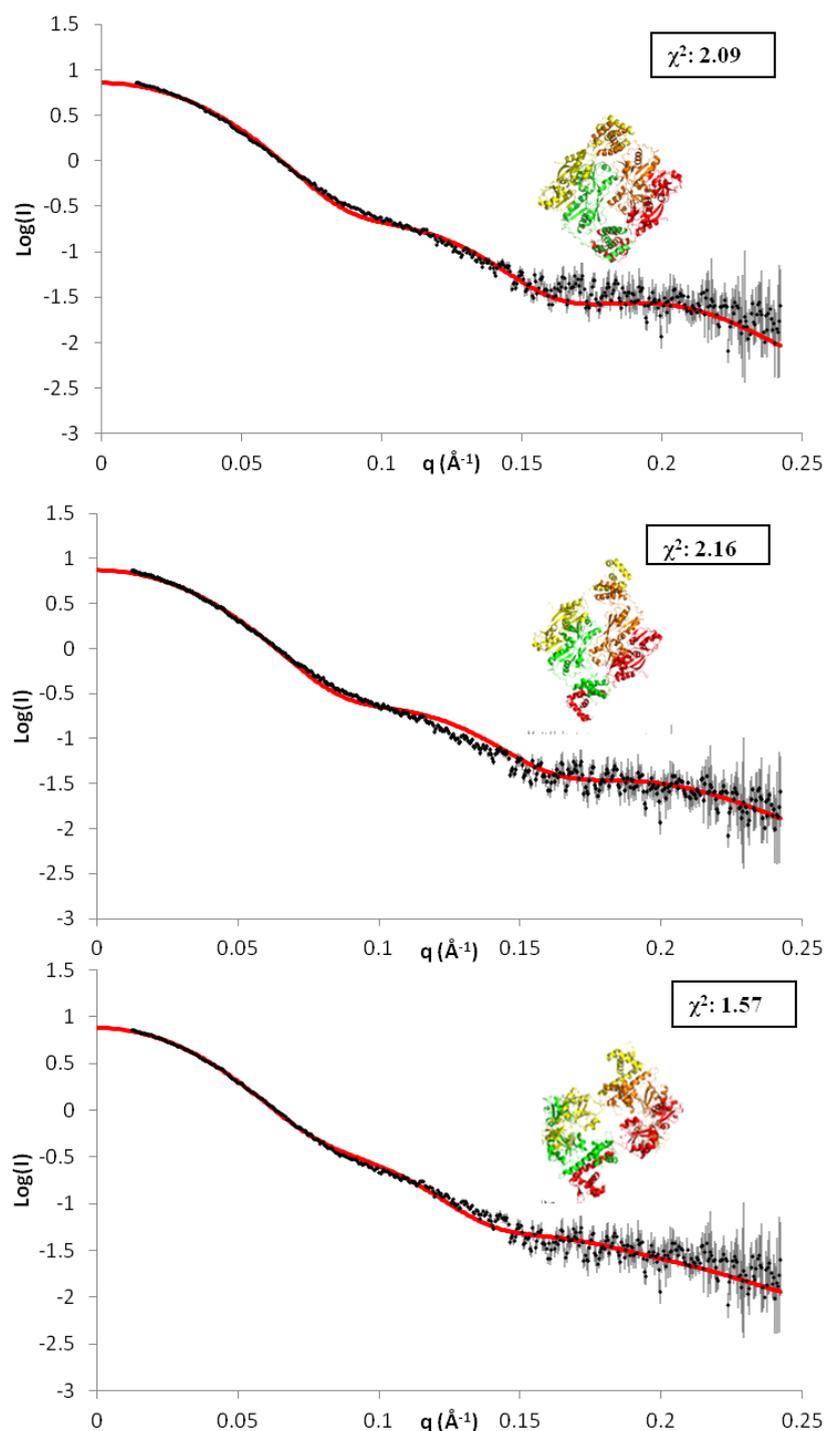
The solution scattering curve obtained in the presence of 100  $\mu$ M sodium salicylate was compared, using CRY SOL, to the calculated scattering curves based on the crystal structure of *apo*-DntR, the TsaR homology model of the open form of a DntR homotetramer and the SASREF-obtained solution structure for *apo*-DntR. These resulting fits (Figure 3.26) make it evident that calculated scattering curves of the crystal and solution structures of *apo*-DntR homotetramers determined during this work, provide poor fits to the experimental data. However, the theoretical curve calculated using the TsaR type homology model of an open-form DntR tetramer shows a much better fit, indicating that the solution structure of salicylate-bound DntR is more similar to this structure than to the compact tetramers seen in either the crystal or solution structures of *apo*-DntR described here. Thus, as was the case for H169TDntR (Section 3.2), the invariant parameters, CRY SOL fits and *ab initio* envelopes derived from SAXS experiments on solutions of salicylate-bound DntR, are all consistent with a DntR tetramer that is closer in conformation to an open TsaR-type tetramer than to the compact tetrameric structure obtained for the solution structure of *apo*-DntR.

### 3.4.2 Rigid body modeling

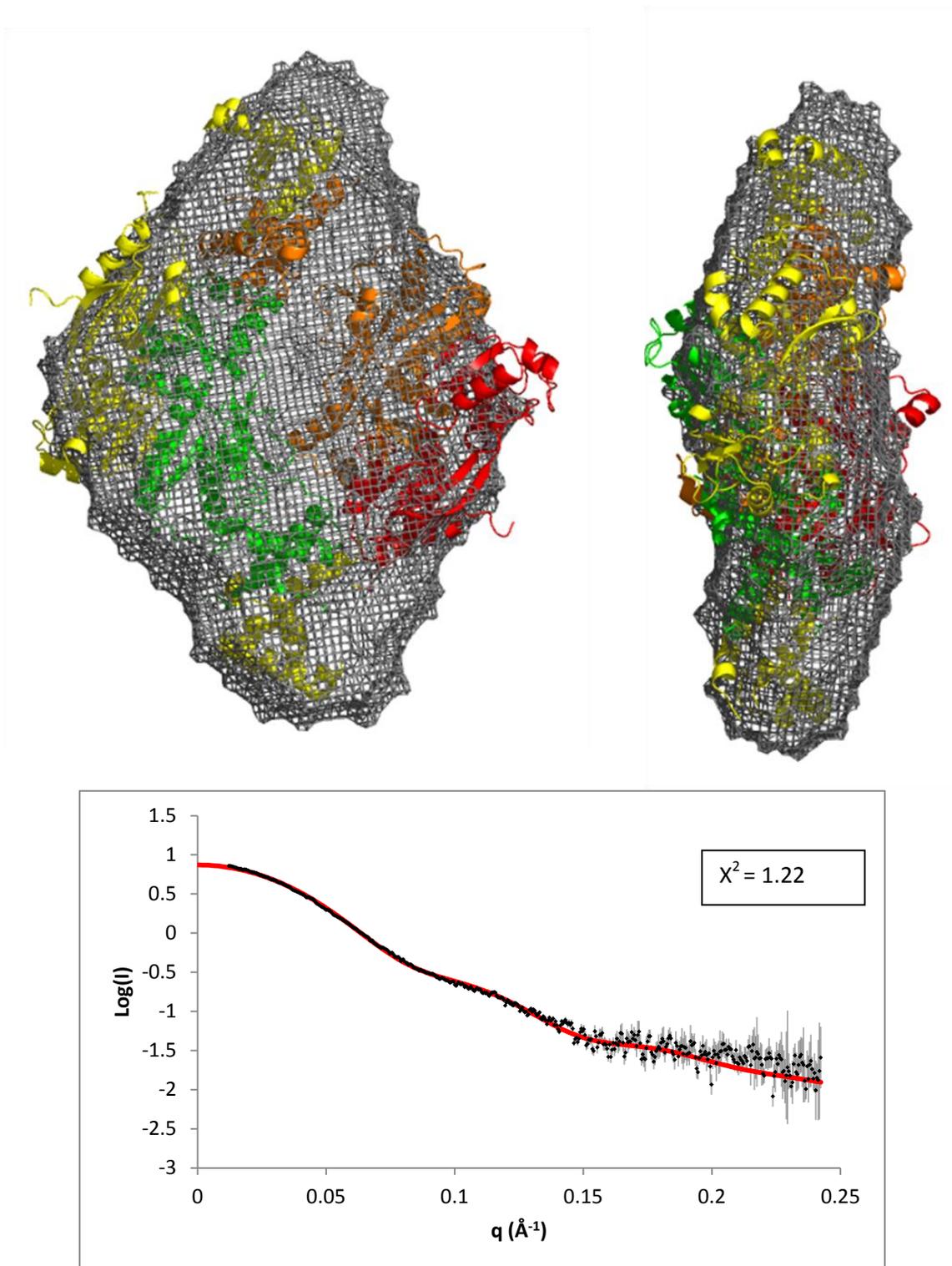
Rigid body modelling of the solution structure of salicylate-bound DntR was performed in SASREF using the same restraints as applied for *apo*-DntR (Section 3.1.2.2). The result shown in Figure 3.21, is consistent with the *ab initio* envelopes obtained and with CRY SOL fits (see above) clearly indicating that in solution *holo*- DntR homotetramers adopt a conformation very reminiscent of the open-form tetramers seen in the crystal structure of TsaR.



**Figure 3.25: SAXS determined *ab initio* molecular envelope of salicylate-bound DntR overlaid with different models of the structure of full length DntR tetramer.** Envelopes were based on 20 independent DAMMIF runs averaged and filtered with DAMAVER, were constructed without any symmetry restrictions and manually overlaid with the Tsar-like open-form DntR homology model (top), the crystal structure of *apo*-DntR (middle) and the solution structure of *apo*-DntR previously generated (bottom).

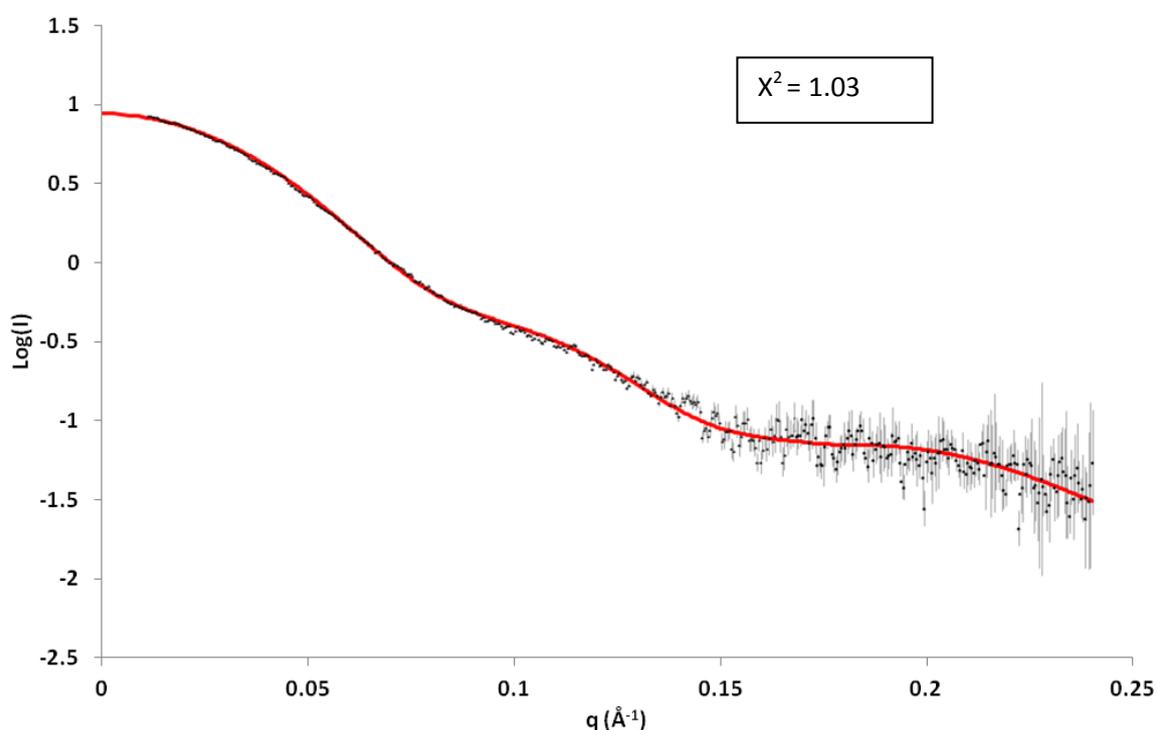


**Figure 3.26: SAXS analysis of a solution of DntR pre-incubated with 100  $\mu\text{M}$  sodium salicylate - CRY SOL fits against the experimental data.** Top: Theoretical scattering curve calculated from solution structure of *apo*-DntR as generated by SASREF (red) vs. the experimental scattering curve obtained from a solution of DntR pre-incubated with 100  $\mu\text{M}$  sodium salicylate (black). Middle: Theoretical scattering curve calculated from the crystal structure of *apo*-DntR (red) vs. the experimental scattering curve obtained from a solution of DntR pre-incubated with 100  $\mu\text{M}$  sodium salicylate (black). Bottom: Theoretical scattering curve calculated from the open-form DntR homology model (red) vs. the experimental scattering curve obtained from a solution of DntR pre-incubated with 100  $\mu\text{M}$  sodium salicylate (black). The resulting  $\chi^2$  values of the fits are shown as inserts.



**Figure 3.27: The SASREF-obtained solution structure of salicylate-bound DntR.** Top SASREF-generated rigid body model of the solution structure of salicylate-bound DntR overlaid with the *ab initio* envelope obtained. Bottom: Theoretical scattering curve calculated from the SASREF-generated model (red) vs. the experimental scattering curve obtained from a solution of DntR pre-incubated with 100  $\mu\text{M}$  sodium salicylate. The  $\chi^2$  value showed as an insert. The poor fit at higher  $q$  range may be caused by local differences between the model obtained and the true solution structure.

As can be seen in Figure 3.27, scattering curves calculated from the SASREF generated solution structure appears to be a good match for the experimental scattering data. However the poor fit at high  $q$  values indicates that there are local differences probably caused by the inflexibility of the rigid bodies used in the SASREF procedure between the structure generated and the actual conformation of the protein in solution. To try to obtain a better fit of the theoretical curve to the experimental data, the IBDs in the SASREF generated solution structure were replaced with those observed for the doubly-salicylate-bound IBDs seen in the crystal structure of *holo*- $\Delta$ N90DntR (PDB 2Y7K, chain A and B) (Devesse et al., 2011). Here the IBCs are expanded to accommodate salicylate and the IBD is slightly closed around the salicylate bound in the IBC. This replacement resulted in a theoretical scattering curve that provides a good fit to the experimental scattering data over the entire  $q$  range (Figure 3.28).



**Figure 3.28: SAXS analysis of salicylate-bound DntR - CRYSOLOG fits against the experimental data.** Theoretical scattering curve calculated from the the SASREF-generated solution structure of salicylate-bound DntR in which the IBDs are closed (red) vs. the experimental scattering curve obtained from a solution of DntR pre-incubated with 100  $\mu$ M sodium salicylate (black). The  $\chi^2$  value is shown as an insert.

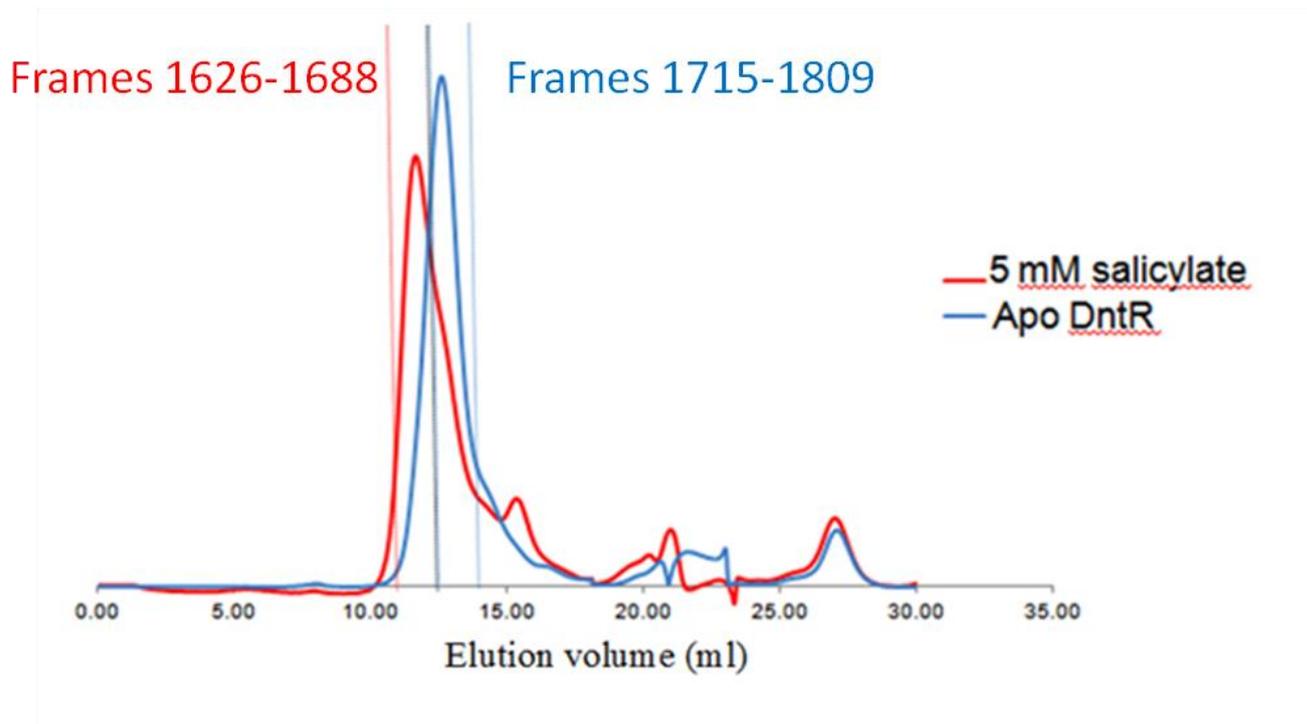
### 3.4.3 Conclusions

The results of the SAXS studies in a solution of salicylate-bound DntR described above are consistent with those observed for the constitutively active H169TDntR and provide further evidence that upon activation DntR homotetramers undergo a conformational change such that they adopt a quaternary structure close to that seen in the crystal structure of Tsar (Monferrer et al., 2010). The solution structure of salicylate-bound DntR thus supports the hypothesis that upon inducer binding, activation of LTRR homotetramers involves a change in conformation from a compact to an open tetrameric configuration. However, given the propensity of LTRR proteins to aggregate (Section 1.3.2), to rule out that the results obtained thus far are artefacts due to interparticle interactions, it was deemed necessary to perform SAXS experiments coupled with online size exclusion chromatography.

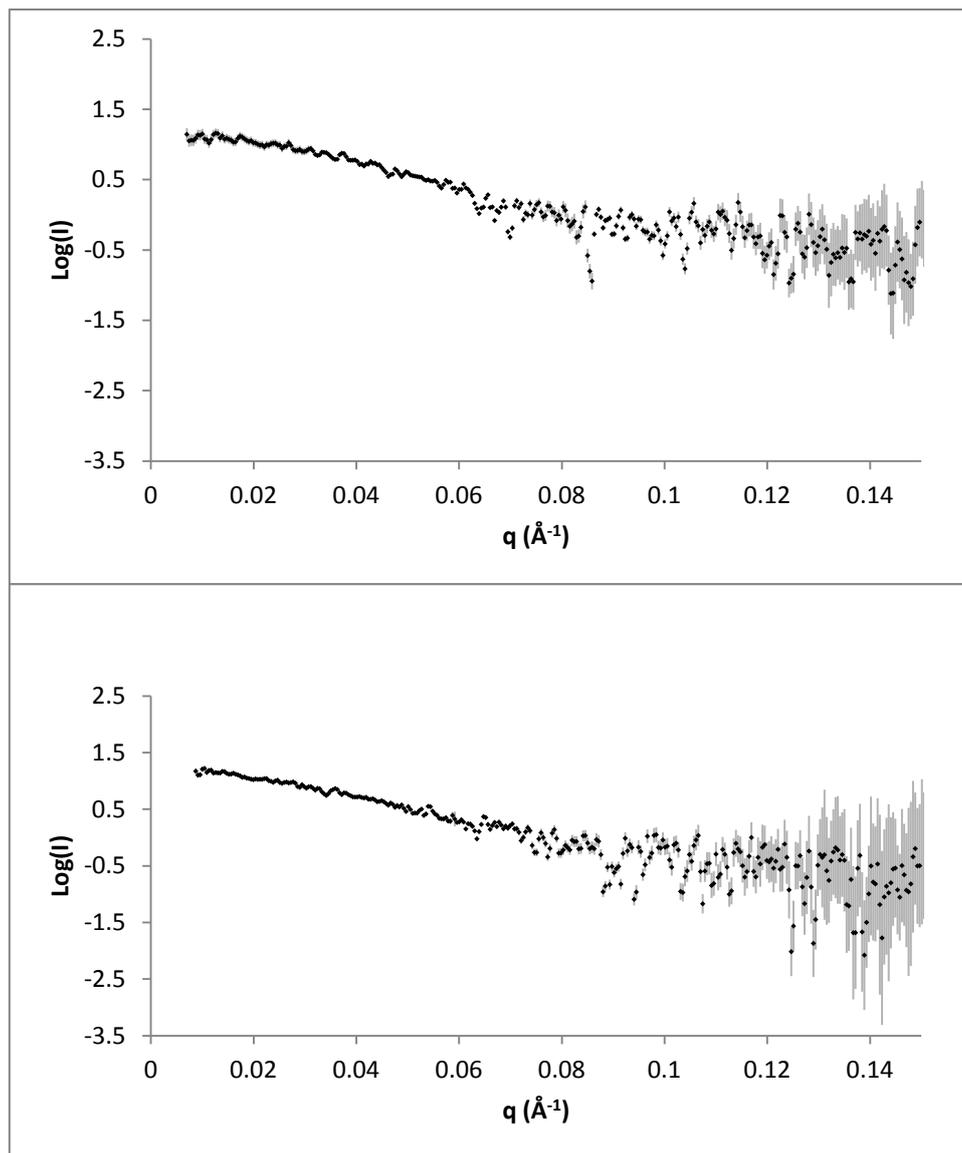
### 3.5 HPLC-coupled SAXS

In order to verify that the observed differences in solution scattering curves and the resulting molecular envelopes and structural models between the *apo*- and *holo*- states of DntR were due to real conformational changes and not caused by small amounts of interparticle interactions in the presence of sodium salicylate, SAXS experiments were repeated, this time coupled with size exclusion chromatography (see Section 2.4.2 for experimental details). The solutions of DntR were applied to a gel filtration column and the resulting eluent exposed directly to the X-ray beam. The elution profiles of *apo*-DntR and DntR pre-incubated with 5 mM sodium salicylate are shown in Figure 3.29 which evidences a clear difference in the behaviour of *apo*- and *holo*-DntR. A total of 3500 frames were collected for each gel filtration run and frames which had near identical  $R_g$  values as determined by AUTOGNOM (Figures 2.19, 2.20 and 3.29), were pooled for manual processing and averaged. The resulting scattering curves are shown in Figure 3.30.

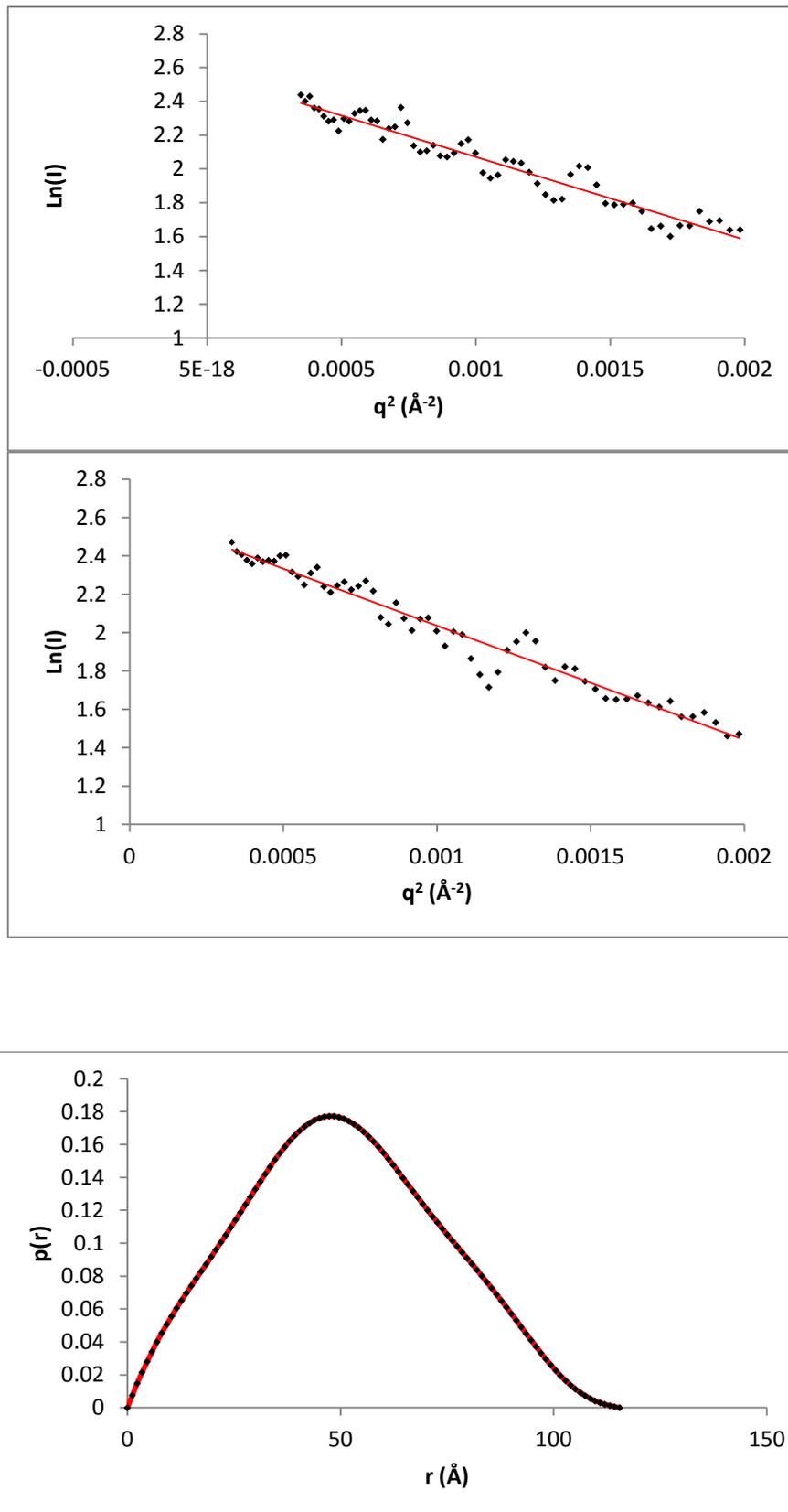
The scattering curves obtained are much noisier than those obtained from the static SAXS measurements described earlier (Section 3.1.2) as the concentration of protein in the eluent is much lower than that used in static SAXS measurements. Nevertheless, it is still possible to obtain satisfactory Guinier regions and  $P(r)$  plots (Figure 3.31) from which the invariant parameters shown in Table 3.7 were derived.



**Figure 3.29: The online gel filtration elution profiles of solutions of *apo*- and *holo*-DntR.** *apo*-DntR (blue) and *apo*-DntR having been pre-incubated with 5 mM sodium salicylate for 30 minutes (red). For the latter 100  $\mu$ M sodium salicylate was also added to the elution buffer. Both runs were conducted with a flow rate of 0.5 ml/min. The position of the frames that were used for manual data processing are shown (red lines for *holo*-DntR, blue lines for *apo*-DntR).



**Figure 3.30: Averaged solution scattering curves for gel filtrated *apo*-(top) and *holo*-DntR (bottom).** Solution scattering curve displaying  $\log(I)$  as a function of  $q$ . To obtain the curves, approximately 80 frames were averaged in both cases. The scattering curves are displayed in the  $q$  range 0 -0.15  $\text{\AA}^{-1}$ .



**Figure 3.31: Guinier plots and  $P(r)$  functions obtained of *apo*-DntR and DntR pre-incubated with 5 mM salicylate.** Guinier regions data points (black) and the resulting fit (red) derived from AUTOGNOM for solutions of *apo*-DntR (Top) and *holo*-DntR (Middle). Bottom:  $P(r)$  functions of *apo*-DntR (blue) and *holo*-DntR (red).

[salicylate] $\mu\text{M}$	$R_g$ ( $\text{\AA}$ )	$I(0)$	$D_{\text{max}}$ ( $\text{\AA}$ )	Porod volume $\text{nm}^3$
0	39.1 (+/-0.5)	4.9 (+/- 0.03)	118	253.3
5000	41.7 (+/-0.5)	5.2 (+/-0.02)	141	255.2

**Table 3.7: Invariant parameters obtained from gel filtrated SAXS scattering curves of DntR in the absence and presence of 5 mM sodium salicylate.** Values were obtained from plots shown in Figure 3.31.

As was the case with static measurements, the invariant parameters indicate that DntR undergoes a conformational change upon the binding of salicylate. Moreover, while the lowered signal to noise ratio and a smaller useable  $q$  range in the HPLC-coupled SAXS experiments cause *ab initio* models obtained from SAXS experiments on solutions subject to gel-filtration immediately before scattering curve measurements to contain fewer features than those obtained in static measurements (data not shown), the values for  $R_g$  and  $D_{\text{max}}$  derived from such experiments are near identical to those obtained in static measurements. This strongly suggests that the structural models derived from higher  $q$ -range static SAXS experiments are valid (i.e. not affected by aggregation), as are conclusions concerning the conformational status of activate and inactive DntR that are based on these.

## 4 Discussion

### 4.1 Summary of results

#### 4.1.1 X-ray crystallography

This work utilized X-ray crystallography to investigate the crystal structures of DntR in its *apo*- state and of the constitutively active mutant H169TDntR.

##### 4.1.1.1 *apo*-DntR

The crystal structure of *apo*-DntR presented in Figure 3.2 is iso-structural to previous DntR crystal structures solved in complex with either thiocyanate or acetate (Smirnova et al., 2004). The electron density of the two N-terminal WTH/linker helix domains of the two DntR monomers in the asymmetric unit is more well defined than in the previously reported DntR crystal structures and allowed for a complete description of the crystal structure of the *apo*-DntR homotetramer. The crystal structure obtained essentially confirms the model of the DntR tetramer previously suggested by (Smirnova et al., 2004). However, the crystal structure of *apo*-DntR also disputes the hypothesis made by Smirnova and colleagues that the ion-bound structures they determined represented DntR homotetramers in an active conformation. Indeed, a comparison of the configuration of the IBCs in the crystal structure for *apo*-DntR obtained here and in the crystal structures of ion-bound DntRs shows that in the latter the IBCs cannot bind salicylate and suggests that the conformation for DntR homotetramers observed by Smirnova and colleagues is actually that of DntR in an inactive state.

##### 4.1.1.2 H169TDntR

The crystals of the constitutively active H169TDntR obtained during this thesis work have very similar unit cell dimensions (Table 3.3) to those obtained for *apo*-DntR (Table 3.1) and the resulting crystal structures are also near identical. However, the IBCs of H169TDntR monomers adopt a much more open conformation, similar to the observed in salicylate-bound  $\Delta$ N90DntR. The ability to induce transcription without a ligand suggests that *apo*-H169TDntR adopts an activated conformation much more readily than does *apo*-DntR. The activated state of DntR has already been proposed to be caused by a closure of the DntR IBD

which may be rendered more favourable by the opening of the IBC as seen in the crystal structure of H169TDntR.

#### 4.1.2 SAXS

To examine the in-solution structures of DntR, free of crystal packing effects, SAXS experiments were performed on solutions of *apo*- and *holo*-DntR and on a solution of H169TDntR.

##### 4.1.2.1 *apo*-DntR

The invariant parameters obtained from solution scattering curves from solutions of *apo*-DntR showed some discrepancy to those expected from the compact homotetrameric crystal structure of *apo*-DntR and suggested that in solution DntR homotetramers in the *apo*- state in solution are less elongated than is observed in its crystal structure. To obtain a proper model of the solution structure of *apo*-DntR homotetramers rigid body modelling was carried out against the experimentally obtained SAXS scattering curve. The resulting model provides a good fit to both the scattering curve and the *ab initio* model envelopes obtained. The resulting solution structure does not greatly change in quaternary structure compared to the crystal structure of the *apo*-DntR homotetramer. However, there is a clear change in the positions of the wHTH dimers flanking the central tetrameric IBD core. These pack more closely to the core than is observed in the crystal structure. A theoretical scattering curve calculated using an open-form DntR homology model based on the crystal structure of TsaR did not provide a good fit to the experimental scattering curve thus excluding the hypothesis that, in solution, *apo*-DntR would adopt an open homotetrameric form.

##### 4.1.2.2 H169TDntR

The invariant parameters obtained from scattering curves measured from solutions of H169TDntR indicated a solution structure that is slightly more elongated than the solution structure of *apo*-DntR. Here, CRY SOL fits and *ab initio* envelopes showed consistently better fits to the open-form DntR homology model based on the crystal structure of TsaR suggesting that the constitutively active H169TDntR has a similar structure. Taken together with the solution structure of *apo*-DntR homotetramers determined during this work (see

above) this strongly supports the idea that upon activation DntR homotetramers change conformation from a compact to a more open form.

#### 4.1.2.3 *holo-DntR*

To confirm the above suggestion, a solution of *holo-DntR* was created through the pre-incubation of *apo-DntR* with 100  $\mu\text{M}$  sodium salicylate. The invariant parameters obtained from the resulting SAXS scattering curves were similar to those obtained for H169TDntR, indicating, in solution, a more elongated quaternary structure for *holo-DntR* than for *apo-DntR*. As was the case for H169TDntR, CRY SOL fits of a theoretical scattering curve calculated using the open-form DntR homology model showed better fit to the experimental data than did scattering curves calculated using either the solution or crystal structures of *apo-DntR* determined during this work.

Rigid body modeling using SASREF resulted in a *holo-DntR* solution structure similar to that of the crystal structure of TsaR in which the LTR homotetramer adopts an elongated conformation with a central cavity between the two IBD dimers in the homotetramer observed. The model of the solution structure of *holo-DntR* homotetramers model could be modified to produce a CRY SOL fit in which  $\chi^2 = 1.03$  by the replacement of the IBD dimers with those observed in the crystal structure of doubly salicylate-bound  $\Delta\text{N90}$  DntR (PDB 2y7k, page 97).

#### 4.1.2.4 *Gel filtrated samples*

Gel-filtration-coupled SAXS experiments were carried out in order to test whether or not the differences in DntR homotetramer conformation observed for *apo-* and *holo-DntR* in static SAXS experiments were, in fact, artefacts caused by a small amount of aggregation upon the addition of 100  $\mu\text{M}$  sodium salicylate to solutions of *apo-DntR*. Table 4.1 shows a reminder of the invariant parameters derived in the various static and HPLC-coupled experiments carried out. As can be seen, the values for these parameters derived from solutions subject to gel-filtration immediately before scattering curve measurements are near identical to those obtained in static measurements. This strongly suggests that the structural models derived from higher  $q$ -range static SAXS experiments are valid (i.e. not affected by

aggregation), as are conclusions concerning the conformational status of activate and inactive DntR that are based on these.

Static Measurements			Gel Filtrated Sample		
[salicylate] $\mu\text{M}$	$R_g$ ( $\text{\AA}$ )	$D_{\text{max}}$ ( $\text{\AA}$ )	[salicylate] $\mu\text{M}$	$R_g$ ( $\text{\AA}$ )	$D_{\text{max}}$ ( $\text{\AA}$ )
0	38.4 (+/-0.6)	118	0	39.2 (+/-0.5)	118
100	41.4 (+/-0.4)	142	5000	41.7 (+/-0.5)	141

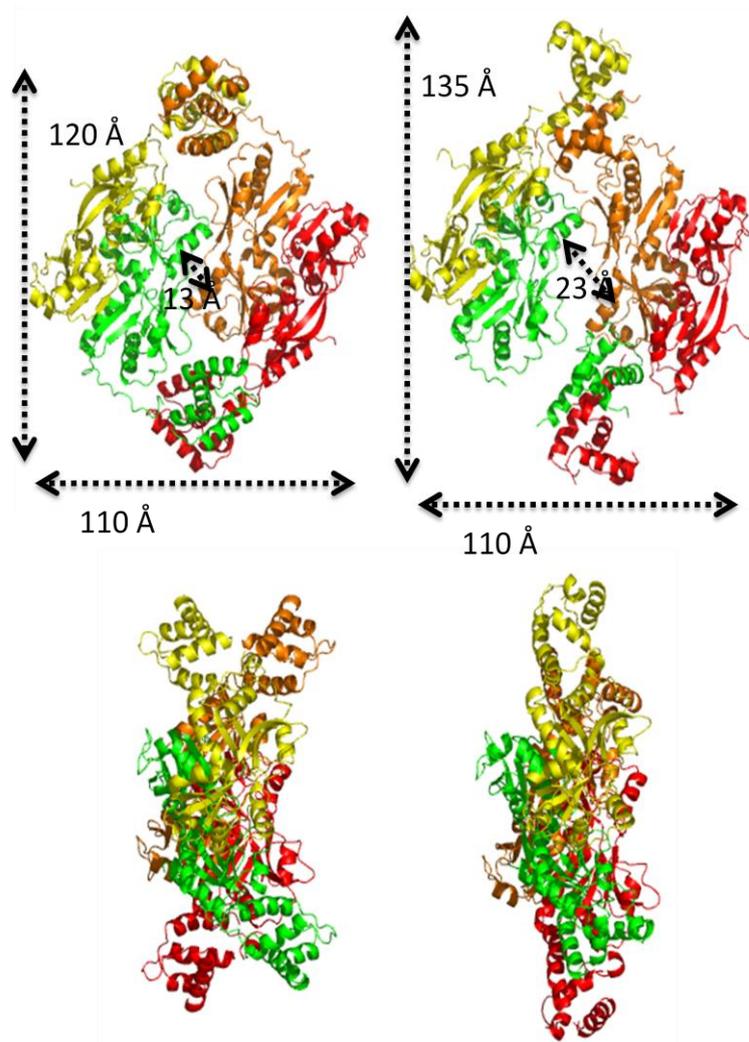
**Table 4.1: invariant SAXS parameters obtained for *apo*- and *holo*-DntR with and without gel filtration prior to measurements.**

## 4.2 The inactive and active states of DntR

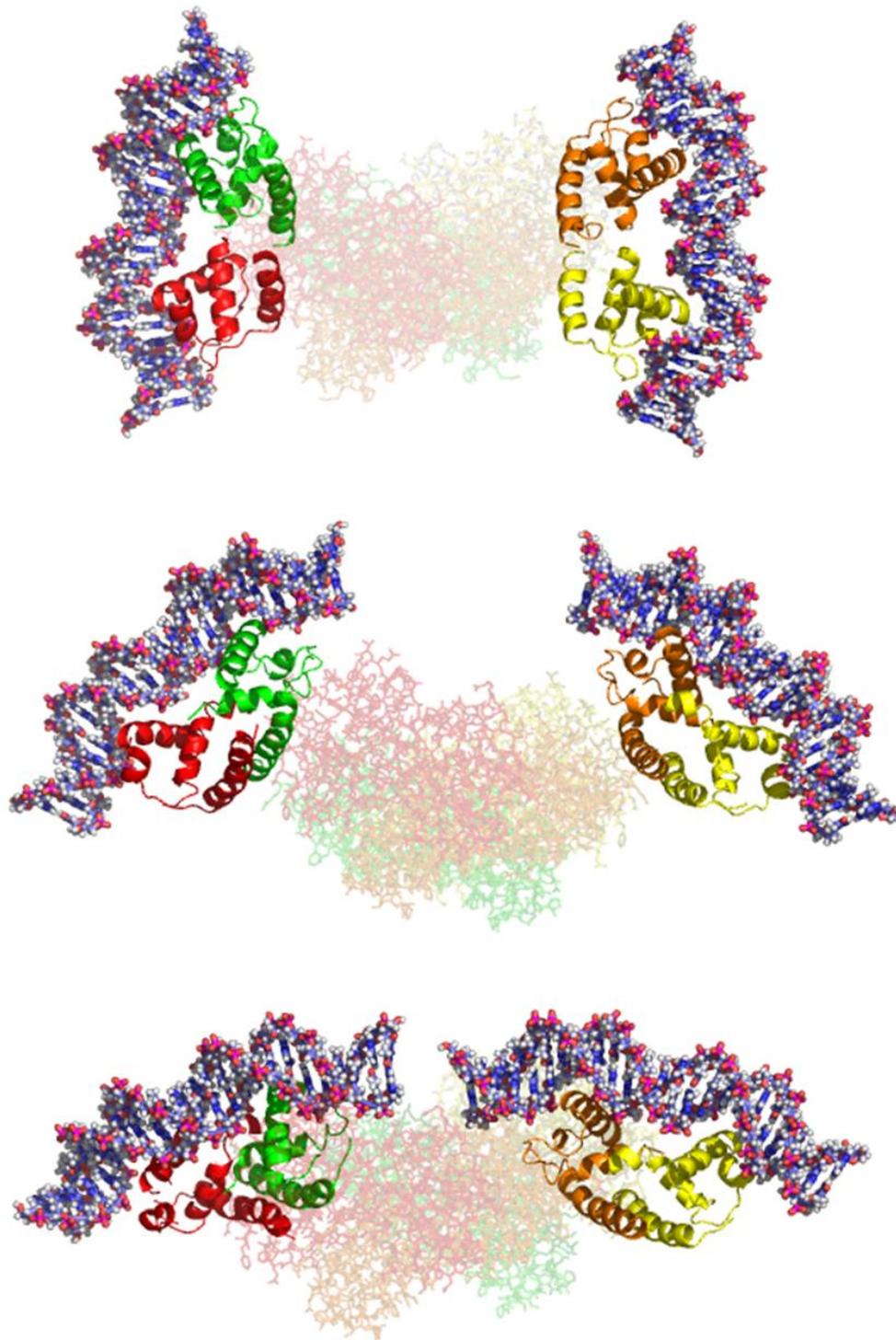
The current consensus is that LTTRs regulate transcription through large conformational changes, which would modify the conformation of bound DNA, upon inducer binding. While previous structural studies (Devesse et al., 2011b; Monferrer et al., 2010b) have suggested that the conformational change in question is a movement from compact to open form homotetramers no LTTR has yet been observed to adopt both configurations. The results of the SAXS studies on solutions of *apo*-DntR, H169TDntR and salicylate-bound DntR here provides the first unequivocal evidence that a LTTR can indeed adopt both conformations (Figure 4.1, left). SAXS studies of *apo*-DntR shows that in solution, inactive DntR homotetramers adopt a conformation that is more compact than observed in the crystal structures of either ion-bound (Smirnova et al. 2004) or *apo*-DntR (this work). Of particular interest in the solution structure of *apo*-DntR homotetramers is the wHTH DBD dimers. These pack much more closely to the tetrameric core than seen in crystal structures and adopt a conformation which would cause the T-N11-A regions bound to these motifs to be almost parallel (Figure 4.2, top).

The SAXS studies reported here for *holo*-DntR and the constitutively active H169TDntR suggest a markedly different conformation for the structure of activated DntR homotetramers. Here the conformation observed for DntR homotetramers is close to that seen in the crystal structure of Tsar. In the open form homotetramers observed the IBD dimers making up the central core of the homotetramer are separated creating a central cavity (Figure 4.1, right) while the wHTH N-terminal domains are displaced such that the bound T-N11-A DNA motifs would adopt an angle to each other of approximately 30 degrees

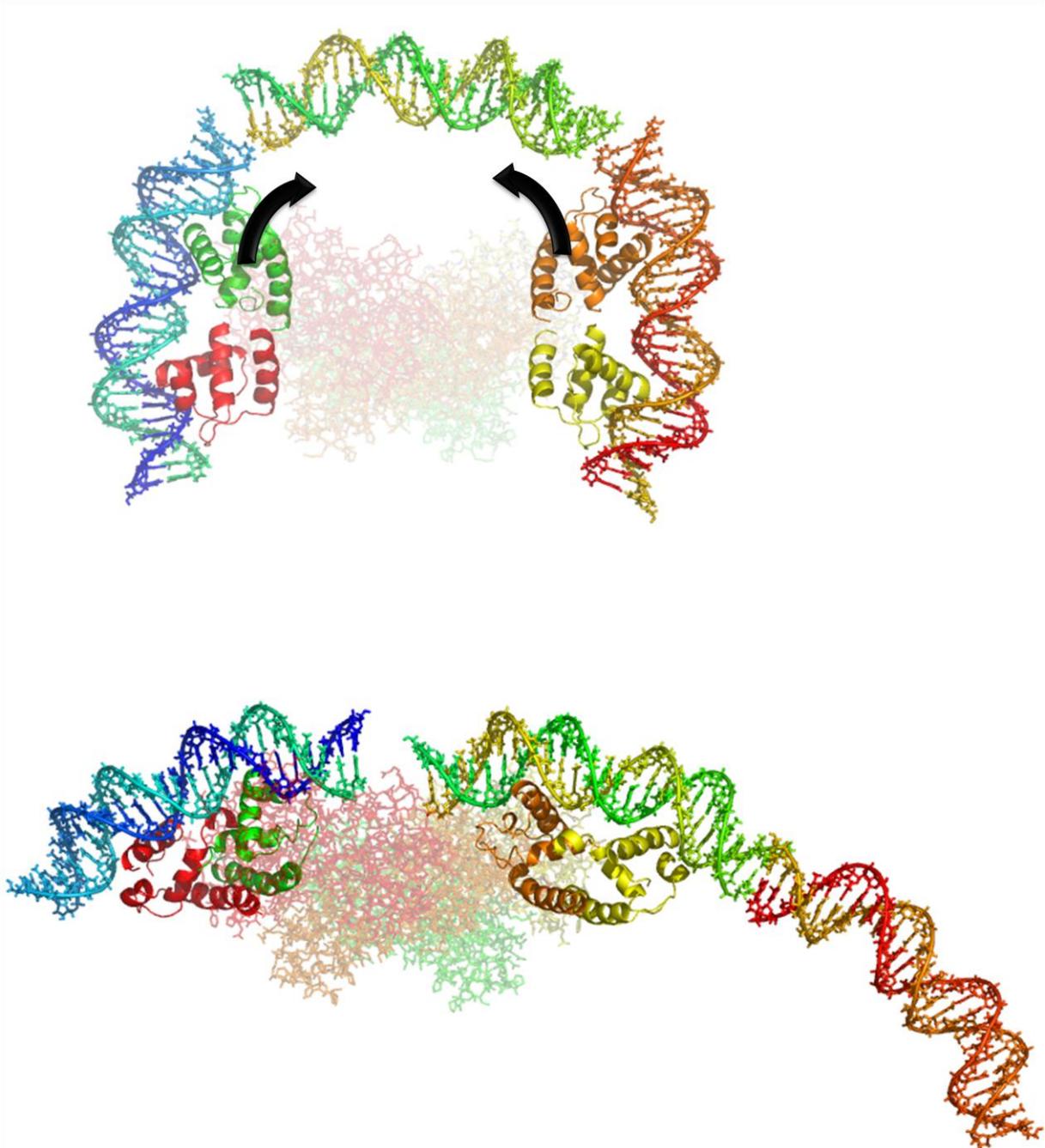
(Figure 4.2, bottom). Intriguingly, comparing the solution structures of *apo*- and *holo*- DntR with the crystal structure of *apo*-DntR it appears that the latter shows a DntR homotetramer in an intermediate state between active and inactive states (Figure 4.2, middle). Without either DNA or inducer molecules bound, the wHTH DBD dimers appear to show some degree of flexibility allowing this intermediate state to occur. It is worth noting that in the models shown in Figure 4.2, a relaxing of the bend in bound promoter region DNA can only occur if the binding site for one (or both) of the DntR DBD dimers is displaced. This would have the effect of pushing the promoter region away from DntR thus making it available for RNA polymerase. A shift such as this was observed for the *clcA* gene as described in Section 1.3.1.



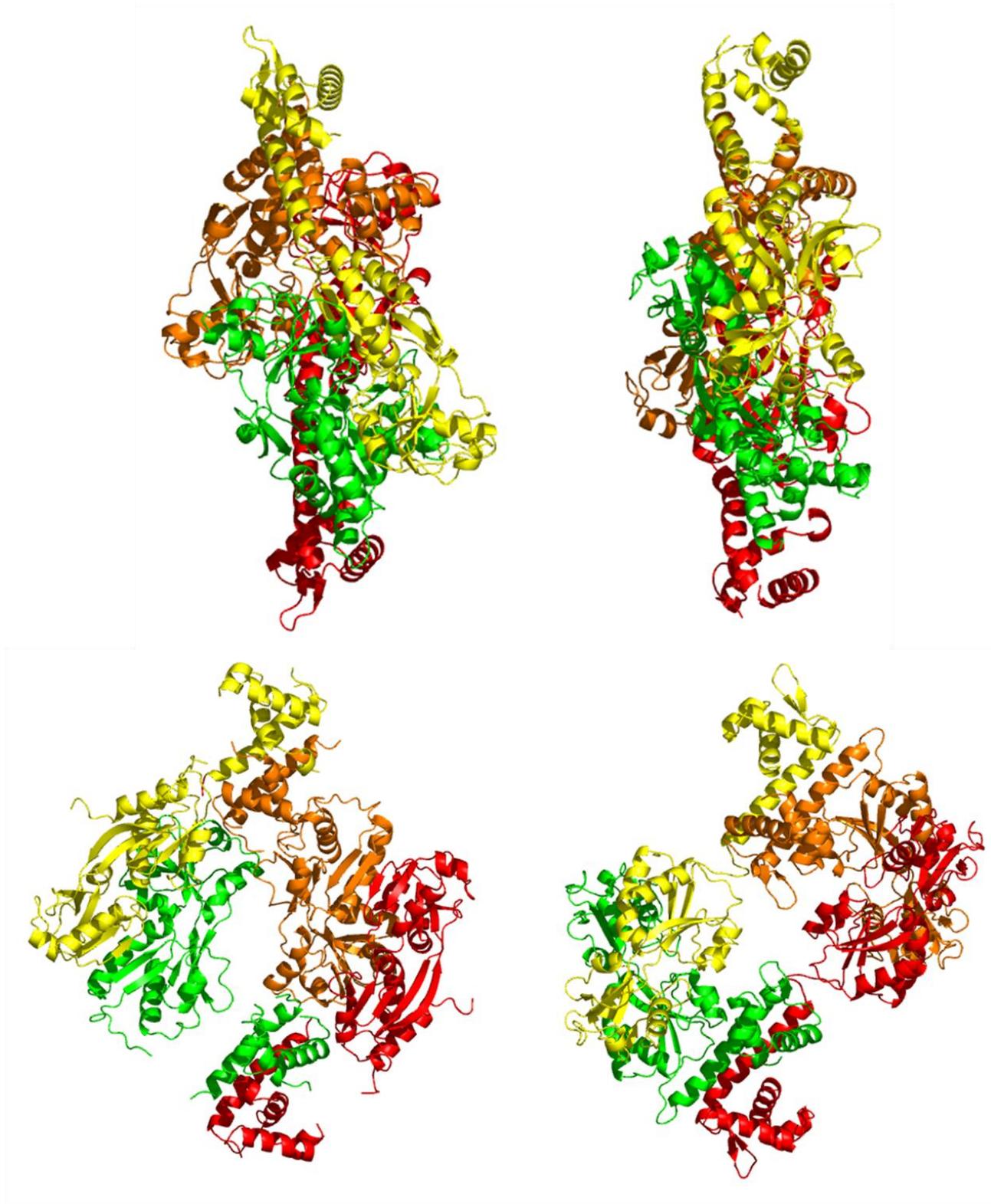
**Figure 4.1: Orthogonal views of the solution structures of *apo*- and *holo*-DntR. SASREF-generated solution structures of *apo*-DntR (left) and *holo*-DntR (right) homotetramers. The distances between the two IBD dimers making up the central core of the homotetramers are depicted as are the overall dimensions of the homotetramers themselves.**



**Figure 4.2: Side views of the models of DNA binding position of the N-terminal wHTH/linker helix domains in the various DntR homotetramers observed.** The position of bound DNA was based on alignments with the crystal structure of the BenM DNA-wHTH/linker helix domain in complex with DNA (Alanazi et al., 2013b). The models shown are: DNA binding to the wHTH domains of the solution structure of *apo*-DntR (top); DNA binding to the wHTH domains as seen in the crystal structure of *apo*-DntR (middle) and DNA binding to the wHTH domains as seen in the solution structure of salicylate-bound DntR (bottom).



**Figure 4.3: The possible displacement of a binding site for bound DNA upon DntR activation.** In order to relax the bend of bound promoter region DNA the binding site for one (or both) of the DntR DBD dimers must be displaced. This would have the effect of pushing the ABS site (red) of the promoter region away from DntR thus making it available for RNA polymerase.



**Figure 4.4:** Comparison of the solution structure of *holo*-DntR and the crystal structure of Tsar. Orthogonal views of the solution structure of salicylate bound DntR (left) and the crystal structure of Tsar (right).

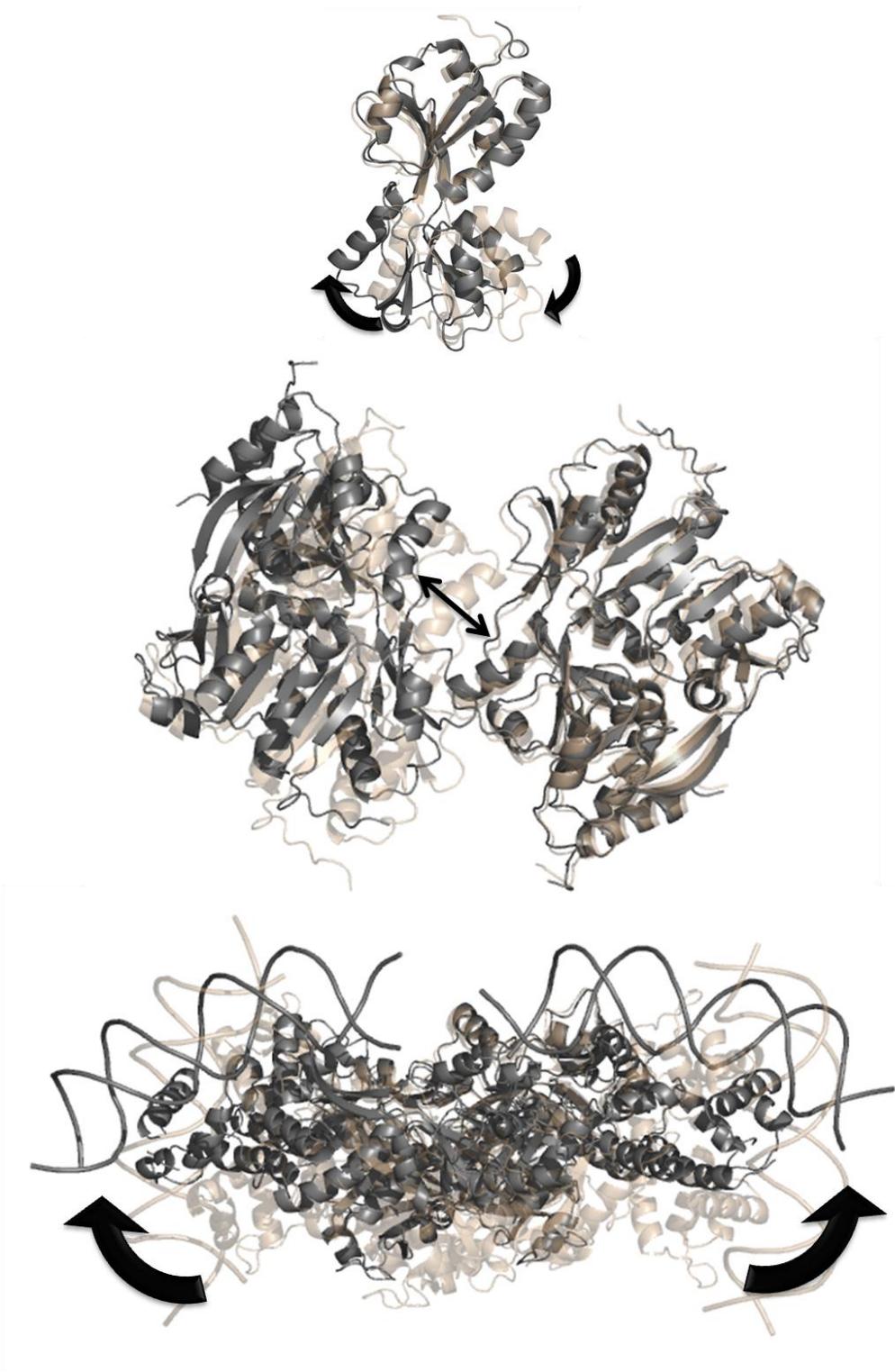
### 4.3 The activation of LTTR proteins

In their recent paper Devesse et al. (Devesse et al., 2011) presented crystal structures of both *apo*- and salicylate-bound  $\Delta N90$ DNTR. For some of the structures presented  $\Delta N90$ DNTR adopted a closed conformation. Comparing this conformation with that of the Inducer Binding Domains observed in the crystal structure of homotetrameric TsaR (Monferrer et al., 2010) provided evidence to support the hypothesis made by Monferrer and co-workers that inactive LTTR tetramers adopted a compact conformation while active, inducer-bound, LTTRs take up a more extended form. However, while the crystal structures of several LTTR proteins have shown that they can adopt either compact or extended conformations, to date no LTTR has been observed to adopt both configurations. Moreover, assignment as to whether a crystal structure represents a particular LTTR in an inactive or active state has been complicated by the fact that the occupancy of the IBCs in LTTR crystal structures does not necessarily correlate with the observation of compact/extended homotetramers.

The work described here is the first time that a LTTR has been observed to adopt both compact and open configurations: A combination of X-ray crystallography and SAXS clearly shows that inactive *apo*-DntR homotetramers adopt a compact conformation while activated *holo*-DntR homotetramers adopt an open configuration very similar to that observed in the crystal structure of TsaR (Figure 4.4). The hypotheses made by (Devesse et al., 2011) and (Monferrer et al., 2010) have thus been confirmed and it is clear that LTTR homotetramers are activated as schematised in Figure 4.5. Here, inactive LTTR homotetramers adopt a compact conformation in which the two head-to-tail C-terminal IBD dimers associate to form a tetrameric core in which the IBCs adopt a conformation that is inconsistent with inducer molecule binding. This central tetrameric core is flanked by two dimeric wHTH DNA binding domains that are parallel with respect to each other, causes bound promoter region DNA to inaccessible for RNA polymerase.

Inducer binding to the IBCs in the central tetrameric IBD core causes these to expand. Allowing a hinge movement of IBD monomers such that they close around the bound inducer molecule. To avoid steric clashes caused by this hinge movement the compact nature of the central IBD core is abolished. This opening of the tetrameric core then causes movement of the dimeric DNA binding domains allows space for the N-terminal wHTH

domains to adopt a flatter interface with bound DNA resulting in a relaxing of the bending which can only occur if the binding site for one (or both) of the DntR DBD dimers is displaced. This would have the effect of pushing the promoter region away from DntR thus making it available for RNA polymerase.



**Figure 4.7: LTR activation as given in the main text.** Inactive DntR (transparent) and active DntR (black)

#### 4.4 Future work.

The work presented in this thesis lays the groundwork for studies of interactions between DntR and DNA. The suggested displacement of the promoter region binding sites should be verified using DNA footprinting studies and could be structurally characterized through the use of small angle neutron scattering (SANS) and atomic force microscopy. Through deuteration/contrast matching SANS would allow the elucidation of the solution structures of DNA-bound *apo*- and *holo*-DntR homotetramers.

To ensure that the results presented here really are representative of all LTTRs it would be necessary to repeat them using a different model protein. As TsaR has already been crystallized in what is suggested to be an activated state but not an inactive state. SAXS experiments on solutions of TsaR would elucidate whether this protein also adopts inactive and active conformations similar to those observed here for DntR.

## Reference list

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 213–221.

Alanazi, A.M., Neidle, E.L., and Momany, C. (2013a). The DNA-binding domain of BenM reveals the structural basis for the recognition of a T-N11-A sequence motif by LysR-type transcriptional regulators. *Acta Crystallogr. D Biol. Crystallogr.* *69*, 1995–2007.

Balleza, E., López-Bojorquez, L.N., Martínez-Antonio, A., Resendis-Antonio, O., Lozada-Chávez, I., Balderas-Martínez, Y.I., Encarnación, S., and Collado-Vides, J. (2009). Regulation by transcription factors in bacteria: beyond description. *FEMS Microbiol. Rev.* *33*, 133–151.

Beck, C.F., and Warren, R.A. (1988). Divergent promoters, a common form of gene organization. *Microbiol. Rev.* *52*, 318–326.

Bourenkov, G.P., and Popov, A.N. (2010). Optimization of data collection taking radiation damage into account. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 409–419.

Brennan, R.G. (1993). The winged-helix DNA-binding motif: another helix-turn-helix takeoff. *Cell* *74*, 773–776.

Brüning, T., Thier, R., and Bolt, H.M. (2002). Nephrotoxicity and nephrocarcinogenicity of dinitrotoluene: new aspects to be considered. *Rev. Environ. Health* *17*, 163–172.

Bundy, B.M., Collier, L.S., Hoover, T.R., and Neidle, E.L. (2002). Synergistic transcriptional activation by one regulatory protein in response to two metabolites. *Proc. Natl. Acad. Sci. U. S. A.* *99*, 7693–7698.

Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K. (1969). Factor stimulating transcription by RNA polymerase. *Nature* *221*, 43–46.

Cases, I., and de Lorenzo, V. (2005). Promoters in the environment: transcriptional regulation in its natural context. *Nat. Rev. Microbiol.* *3*, 105–118.

Cases, I., de Lorenzo, V., and Ouzounis, C.A. (2003). Transcription regulation and environmental adaptation in bacteria. *Trends Microbiol.* *11*, 248–253.

Choi, H., Kim, S., Mukhopadhyay, P., Cho, S., Woo, J., Storz, G., and Ryu, S.E. (2001). Structural basis of the redox switch in the OxyR transcription factor. *Cell* *105*, 103–113.

Coco, W.M., Rothmel, R.K., Henikoff, S., and Chakrabarty, A.M. (1993). Nucleotide sequence and initial functional characterization of the *clcR* gene encoding a LysR family activator of the *clcABD* chlorocatechol operon in *Pseudomonas putida*. *J. Bacteriol.* *175*, 417–427.

Collado-Vides, J., Magasanik, B., and Gralla, J.D. (1991). Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol. Rev.* *55*, 371–394.

- Collier, L.S., Gaines, G.L., 3rd, and Neidle, E.L. (1998). Regulation of benzoate degradation in *Acinetobacter* sp. strain ADP1 by BenM, a LysR-type transcriptional activator. *J. Bacteriol.* *180*, 2493–2501.
- Colyer, T.E., and Kredich, N.M. (1994). Residue threonine-149 of the *Salmonella typhimurium* CysB transcription activator: mutations causing constitutive expression of positively regulated genes of the cysteine regulon. *Mol. Microbiol.* *13*, 797–805.
- Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* *292*, 1863–1876.
- Craven, S.H., Ezezika, O.C., Haddad, S., Hall, R.A., Momany, C., and Neidle, E.L. (2009). Inducer responses of BenM, a LysR-type transcriptional regulator from *Acinetobacter baylyi* ADP1. *Mol. Microbiol.* *72*, 881–894.
- Devesse, L., Smirnova, I., Lönneborg, R., Kapp, U., Brzezinski, P., Leonard, G.A., and Dian, C. (2011). Crystal structures of DntR inducer binding domains in complex with salicylate offer insights into the activation of LysR-type transcriptional regulators. *Mol. Microbiol.* *81*, 354–367.
- Dickerson, R.E., Goodsell, D.S., and Neidle, S. (1994). ...the tyranny of the lattice... *Proc. Natl. Acad. Sci. U. S. A.* *91*, 3579–3583.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 486–501.
- Evans, P.R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* *67*, 282–292.
- Ezezika, O.C., Haddad, S., Clark, T.J., Neidle, E.L., and Momany, C. (2007a). Distinct effector-binding sites enable synergistic transcriptional activation by BenM, a LysR-type regulator. *J. Mol. Biol.* *367*, 616–629.
- Ezezika, O.C., Haddad, S., Neidle, E.L., and Momany, C. (2007b). Oligomerization of BenM, a LysR-type transcriptional regulator: structural basis for the aggregation of proteins in this family. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* *63*, 361–368.
- Fisher, R.F., and Long, S.R. (1993). Interactions of NodD at the nod Box: NodD binds to two distinct sites on the same face of the helix and induces a bend in the DNA. *J. Mol. Biol.* *233*, 336–348.
- Franke, D., and Svergun, D.I. (2009). DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Cryst* *42*, 342–346.
- French, S., and Wilson, K. (1978). On the treatment of negative intensity observations. *Acta Cryst* *34*, 517–525.
- Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* *292*, 1876–1882.

- Goulian, M. (2004). Robust control in bacterial regulatory circuits. *Curr. Opin. Microbiol.* 7, 198–202.
- Henikoff, S., Haughn, G.W., Calvo, J.M., and Wallace, J.C. (1988). A large family of bacterial activator proteins. *Proc. Natl. Acad. Sci. U. S. A.* 85, 6602–6606.
- De las Heras, A., Chavarría, M., and de Lorenzo, V. (2011). Association of *dnt* genes of *Burkholderia* sp. DNT with the substrate-blind regulator DntR draws the evolutionary itinerary of 2,4-dinitrotoluene biodegradation. *Mol. Microbiol.* 82, 287–299.
- Hryniewicz, M.M., and Kredich, N.M. (1995). Hydroxyl radical footprints and half-site arrangements of binding sites for the CysB transcriptional activator of *Salmonella typhimurium*. *J. Bacteriol.* 177, 2343–2353.
- Huffman, J.L., and Brennan, R.G. (2002). Prokaryotic transcription regulators: more than just the helix-turn-helix motif. *Curr. Opin. Struct. Biol.* 12, 98–106.
- Incardona, M.-F., Bourenkov, G.P., Levik, K., Pieritz, R.A., Popov, A.N., and Svensson, O. (2009). EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. *J. Synchrotron Radiat.* 16, 872–879.
- Jacob, F., and Monod, J. (1959). Genes of structure and genes of regulation in the biosynthesis of proteins. *Comptes Rendus Hebd. Séances Académie Sci.* 249, 1282–1284.
- Johnson, G.R., Jain, R.K., and Spain, J.C. (2002). Origins of the 2,4-dinitrotoluene pathway. *J. Bacteriol.* 184, 4219–4232.
- Jourdan, A.D., and Stauffer, G.V. (1998). Mutational analysis of the transcriptional regulator GcvA: amino acids important for activation, repression, and DNA binding. *J. Bacteriol.* 180, 4865–4871.
- Ju, M.-J., Yang, D.-H., Takahara, N., Hayashi, K., Toko, K., Lee, S.-W., and Kunitake, T. (2007). Landmine detection: improved binding of 2,4-dinitrotoluene in a gamma-CD/metal oxide matrix and its sensitive detection via a cyclic surface polarization impedance (cSPI) method. *Chem. Commun. Camb. Engl.* 2630–2632.
- Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* 66, 133–144.
- Van Keulen, G., Girbal, L., van den Bergh, E.R., Dijkhuizen, L., and Meijer, W.G. (1998). The LysR-type transcriptional regulator CbbR controlling autotrophic CO<sub>2</sub> fixation by *Xanthobacter flavus* is an NADPH sensor. *J. Bacteriol.* 180, 1411–1417.
- Kivisaar, M. (2009). Degradation of nitroaromatic compounds: a model to study evolution of metabolic pathways. *Mol. Microbiol.* 74, 777–781.
- Koch, M.H., Vachette, P., and Svergun, D.I. (2003). Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* 36, 147–227.

- Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* *36*, 1277–1282.
- Kullik, I., Stevens, J., Toledano, M.B., and Storz, G. (1995). Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for DNA binding and multimerization. *J. Bacteriol.* *177*, 1285–1291.
- Lahiri, A., Das, P., and Chakravorty, D. (2009). Salmonella Typhimurium: insight into the multi-faceted role of the LysR-type transcriptional regulators in Salmonella. *Int. J. Biochem. Cell Biol.* *41*, 2129–2133.
- Lönneborg, R., Smirnova, I., Dian, C., Leonard, G.A., and Brzezinski, P. (2007). In Vivo and in Vitro Investigation of Transcriptional Regulation by DntR. *J. Mol. Biol.* *372*, 571–582.
- Lönneborg, R., Varga, E., and Brzezinski, P. (2012). Directed evolution of the transcriptional regulator DntR: isolation of mutants with improved DNT-response. *PLoS One* *7*, e29994.
- Lozada-Chávez, I., Angarica, V.E., Collado-Vides, J., and Contreras-Moreira, B. (2008). The role of DNA-binding specificity in the evolution of bacterial regulatory networks. *J. Mol. Biol.* *379*, 627–643.
- Madan Babu, M., and Teichmann, S.A. (2003). Functional determinants of transcription factors in *Escherichia coli*: protein families and binding sites. *Trends Genet. TIG* *19*, 75–79.
- Maddocks, S.E., and Oyston, P.C.F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* *154*, 3609–3623.
- Martínez-Antonio, A., Janga, S.C., Salgado, H., and Collado-Vides, J. (2006). Internal-sensing machinery directs the activity of the regulatory network in *Escherichia coli*. *Trends Microbiol.* *14*, 22–27.
- Mascher, T. (2006). Intramembrane-sensing histidine kinases: a new family of cell envelope stress sensors in Firmicutes bacteria. *FEMS Microbiol. Lett.* *264*, 133–144.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* *40*, 658–674.
- McFall, S.M., Klem, T.J., Fujita, N., Ishihama, A., and Chakrabarty, A.M. (1997). DNase I footprinting, DNA bending and in vitro transcription analyses of ClcR and CatR interactions with the clcABD promoter: evidence of a conserved transcriptional activation mechanism. *Mol. Microbiol.* *24*, 965–976.
- McKay, D.B., and Steitz, T.A. (1981). Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature* *290*, 744–749.
- Van der Meer, J.R., Frijters, A.C., Leveau, J.H., Eggen, R.I., Zehnder, A.J., and de Vos, W.M. (1991). Characterization of the *Pseudomonas* sp. strain P51 gene tcbR, a LysR-type transcriptional activator of the tcbCDEF chlorocatechol oxidative operon, and analysis of the regulatory region. *J. Bacteriol.* *173*, 3700–3708.

- Mertens, H.D.T., and Svergun, D.I. (2010). Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* *172*, 128–141.
- Momany, C., and Neidle, E.L. (2012). Defying stereotypes: the elusive search for a universal model of LysR-type regulation. *Mol. Microbiol.* *83*, 453–456.
- Monferrer, D., Tralau, T., Kertesz, M.A., Dix, I., Solà, M., and Usón, I. (2010). Structural studies on the full-length LysR-type regulator TsaR from *Comamonas testosteroni* T-2 reveal a novel open conformation of the tetrameric LTTR fold. *Mol. Microbiol.* *75*, 1199–1214.
- Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K., and Senda, T. (2003). Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. *J. Mol. Biol.* *328*, 555–566.
- Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC 5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* *67*, 355–367.
- Nurizzo, D., Mairs, T., Guijarro, M., Rey, V., Meyer, J., Fajardo, P., Chavanne, J., Biasci, J.-C., McSweeney, S., and Mitchell, E. (2006). The ID23-1 structural biology beamline at the ESRF. *J. Synchrotron Radiat.* *13*, 227–238.
- Pareja, E., Pareja-Tobes, P., Manrique, M., Pareja-Tobes, E., Bonal, J., and Tobes, R. (2006). ExtraTrain: a database of Extragenic regions and Transcriptional information in prokaryotic organisms. *BMC Microbiol.* *6*, 29.
- Parsek, M.R., Kivisaar, M., and Chakrabarty, A.M. (1995). Differential DNA bending introduced by the *Pseudomonas putida* LysR-type regulator, CatR, at the plasmid-borne pheBA and chromosomal catBC promoters. *Mol. Microbiol.* *15*, 819–828.
- Pérez-Martín, J., Rojo, F., and de Lorenzo, V. (1994). Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. *Microbiol. Rev.* *58*, 268–290.
- Pernot, P., Round, A., Barrett, R., De Maria Antolinos, A., Gobbo, A., Gordon, E., Huet, J., Kieffer, J., Lentini, M., Mattenet, M., et al. (2013). Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* *20*, 660–664.
- Petoukhov, M.V., and Svergun, D.I. (2005). Global rigid body modelling of macromolecular complexes against small-angle scattering data. *Biophys J* *89*, 1237–1250.
- Piñeiro, S., Olekhovich, I., and Gussin, G.N. (1997). DNA bending by the TrpI protein of *Pseudomonas aeruginosa*. *J. Bacteriol.* *179*, 5407–5413.
- Ptashne, M., and Gann, A. (2001). Transcription initiation: imposing specificity by localization. *Essays Biochem.* *37*, 1–15.
- Ptashne, M. (1965). The detachment and maturation of conserved lambda prophage DNA. *J. Mol. Biol.* *11*, 90–96.

- Putnam, C.D., Hammel, M., Hura, G.L., and Tainer, J.A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* *40*, 191–285.
- Quade, N., Dieckmann, M., Haffke, M., Heroven, A.K., Dersch, P., and Heinz, D.W. (2011). Structure of the effector-binding domain of the LysR-type transcription factor RovM from *Yersinia pseudotuberculosis*. *Acta Crystallogr. D Biol. Crystallogr.* *67*, 81–90.
- Reen, F.J., Barret, M., Fargier, E., O’Muinneacháin, M., and O’Gara, F. (2013a). Molecular evolution of LysR-type transcriptional regulation in *Pseudomonas aeruginosa*. *Mol. Phylogenet. Evol.* *66*, 1041–1049.
- Reen, F.J., Haynes, J.M., Mooij, M.J., and O’Gara, F. (2013b). A Non-Classical LysR-Type Transcriptional Regulator PA2206 Is Required for an Effective Oxidative Stress Response in *Pseudomonas aeruginosa*. *PLoS ONE* *8*, e54479.
- Ruangprasert, A., Craven, S.H., Neidle, E.L., and Momany, C. (2010). Full-length structures of BenM and two variants reveal different oligomerization schemes for LysR-type transcriptional regulators. *J. Mol. Biol.* *404*, 568–586.
- Sainsbury, S., Lane, L.A., Ren, J., Gilbert, R.J., Saunders, N.J., Robinson, C.V., Stuart, D.I., and Owens, R.J. (2009). The structure of CrgA from *Neisseria meningitidis* reveals a new octameric assembly state for LysR transcriptional regulators. *Nucleic Acids Res.* *37*, 4545–4558.
- Sainsbury, S., Ren, J., Nettleship, J.E., Saunders, N.J., Stuart, D.I., and Owens, R.J. (2010). The structure of a reduced form of OxyR from *Neisseria meningitidis*. *BMC Struct. Biol.* *10*, 10.
- Sainsbury, S., Ren, J., Saunders, N.J., Stuart, D.I., and Owens, R.J. (2012). Structure of the regulatory domain of the LysR family regulator NMB2055 (MetR-like protein) from *Neisseria meningitidis*. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* *68*, 730–737.
- De Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., Gobbo, A., Guijarro, M., Lentini, M., Lavault, B., et al. (2012). ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *J. Synchrotron Radiat.* *19*, 455–461.
- Schell, M.A. (1993). Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* *47*, 597–626.
- Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* *31*, 3381–3385.
- Smirnova, I.A., Dian, C., Leonard, G.A., McSweeney, S., Birse, D., and Brzezinski, P. (2004). Development of a Bacterial Biosensor for Nitrotoluenes: The Crystal Structure of the Transcriptional Regulator DntR. *J. Mol. Biol.* *340*, 405–418.
- Stevens, A., and Rhoton, J.C. (1975). Characterization of an inhibitor causing potassium chloride sensitivity of an RNA polymerase from T4 phage-infected *Escherichia coli*. *Biochemistry (Mosc.)* *14*, 5074–5079.

- Stragier, P., and Patte, J.C. (1983). Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli*. III. Nucleotide sequence and regulation of the *lysR* gene. *J. Mol. Biol.* *168*, 333–350.
- Suen, W.C., and Spain, J.C. (1993). Cloning and characterization of *Pseudomonas* sp. strain DNT genes for 2,4-dinitrotoluene degradation. *J. Bacteriol.* *175*, 1831–1837.
- Svergun, D.I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* *76*, 2879–2886.
- Svergun, D.I., Barberato, C., and Koch, M.H.J. (1995). CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J Appl Cryst* *28*, 768–773.
- Svintradze, D.V., Peterson, D.L., Collazo-Santiago, E.A., Lewis, J.P., and Wright, H.T. (2013). Structures of the *Porphyromonas gingivalis* OxyR regulatory domain explain differences in expression of the OxyR regulon in *Escherichia coli* and *P. gingivalis*. *Acta Crystallogr. D Biol. Crystallogr.* *69*, 2091–2103.
- Taylor, J.L., De Silva, R.S., Kovacicova, G., Lin, W., Taylor, R.K., Skorupski, K., and Kull, F.J. (2012). The crystal structure of AphB, a virulence gene activator from *Vibrio cholerae*, reveals residues that influence its response to oxygen and pH. *Mol. Microbiol.* *83*, 457–470.
- Toledano, M.B., Kullik, I., Trinh, F., Baird, P.T., Schneider, T.D., and Storz, G. (1994). Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: a mechanism for differential promoter selection. *Cell* *78*, 897–909.
- Tralau, T., Mampel, J., Cook, A.M., and Ruff, J. (2003). Characterization of TsaR, an oxygen-sensitive LysR-type regulator for the degradation of p-toluenesulfonate in *Comamonas testosteroni* T-2. *Appl. Environ. Microbiol.* *69*, 2298–2305.
- Tropel, D., and van der Meer, J.R. (2004). Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiol. Mol. Biol. Rev. MMBR* *68*, 474–500, Table of contents.
- Tyrrell, R., Verschueren, K.H., Dodson, E.J., Murshudov, G.N., Addy, C., and Wilkinson, A.J. (1997). The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. *Struct. Lond. Engl.* *1993* *5*, 1017–1032.
- Ulrich, L.E., Koonin, E.V., and Zhulin, I.B. (2005). One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol.* *13*, 52–56.
- Volkov, V.V., and Svergun, D.I. (2003). Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* *36*, 860–864.
- Wang, L., Helmann, J.D., and Winans, S.C. (1992). The *A. tumefaciens* transcriptional activator OccR causes a bend at a target promoter, which is partially relaxed by a plant tumor metabolite. *Cell* *69*, 659–667.
- Weber, I.T., and Steitz, T.A. (1987). Structure of a complex of catabolite gene activator protein and cyclic AMP refined at 2.5 Å resolution. *J. Mol. Biol.* *198*, 311–326.

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.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* *67*, 235–242.

Xu, N., Yu, S., Moniot, S., Weyand, M., and Blankenfeldt, W. (2012). Crystallization and preliminary crystal structure analysis of the ligand-binding domain of PqsR (MvfR), the *Pseudomonas* quinolone signal (PQS) responsive quorum-sensing transcription factor of *Pseudomonas aeruginosa*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* *68*, 1034–1039.

Xu, Y., Heath, R.J., Li, Z., Rock, C.O., and White, S.W. (2001). The FadR-DNA complex. Transcriptional control of fatty acid metabolism in *Escherichia coli*. *J. Biol. Chem.* *276*, 17373–17379.

Zerrad, L., Merli, A., Schröder, G.F., Varga, A., Gráczér, É., Pernot, P., Round, A., Vas, M., and Bowler, M.W. (2011). A spring-loaded release mechanism regulates domain movement and catalysis in phosphoglycerate kinase. *J. Biol. Chem.* *286*, 14040–14048.

Zhou, X., Lou, Z., Fu, S., Yang, A., Shen, H., Li, Z., Feng, Y., Bartlam, M., Wang, H., and Rao, Z. (2010). Crystal structure of ArgP from *Mycobacterium tuberculosis* confirms two distinct conformations of full-length LysR transcriptional regulators and reveals its function in DNA binding and transcriptional regulation. *J. Mol. Biol.* *396*, 1012–1024.

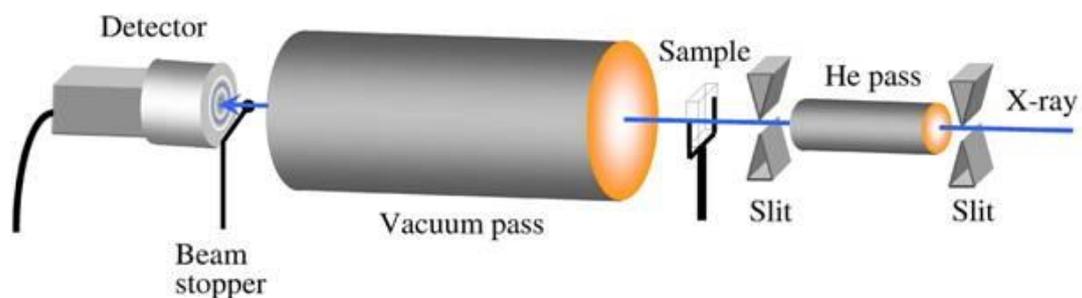
## 5. Appendix: Small Angle X-Ray Scattering

The following section was partially adapted from master's thesis "Insulin fibrillation inhibited with the flavonoids baicalein and epigallocatechin gallate characterized by Thioflavin T assays and Small Angle X-ray Scattering" (Lerche, M, University of Copenhagen, Denmark, 2010).

### 5.1 Principles

Small Angle X-Ray Scattering (SAXS) (Figure 5.1) is a powerful structural biology technique. Unlike X-ray crystallography it does not require the growth of crystals and can thus be utilized to investigate the structure of proteins in solution. SAXS can be used to investigate very large complexes, with molecular weights in the mDa range (Putnam et al., 2007).

While there is no theoretical limit on the resolution that can be obtained for solution structures using SAXS, it is still only possible to produce low resolution models in the 5-10 Å range (Koch et al., 2003; Mertens and Svergun, 2010).



**Figure 5.1: Setup of a SAXS experiment.** The X-rays are scattered by the sample and the scattering is recorded on a 2-D detector. Image from <http://www.saga-ls.jp/images/image/English/Users%20Guide/Beamlines/BL15>

While SAXS and X-ray crystallography have a number of shared characteristics they are fundamentally different. X-ray crystallography measures the diffraction of particles in an ordered lattice in a crystal and then determines the electron density (Putnam et al., 2007). SAXS relies on contrast variation and the scattering signal is derived from the difference of electron density between the disordered macromolecules of interest and their solvent (Koch et al., 2003). This contrast is quite small and its measurements puts severe practical limitations on the resolution of SAXS data (Koch et al., 2003).

### 5.1.1 Scattering intensities

The Scattering pattern of a solution containing randomly oriented particles is a function of  $q$  which is defined as:

$$q = \frac{4\pi \sin(\theta)}{\lambda} \quad (\text{Equation 5.1})$$

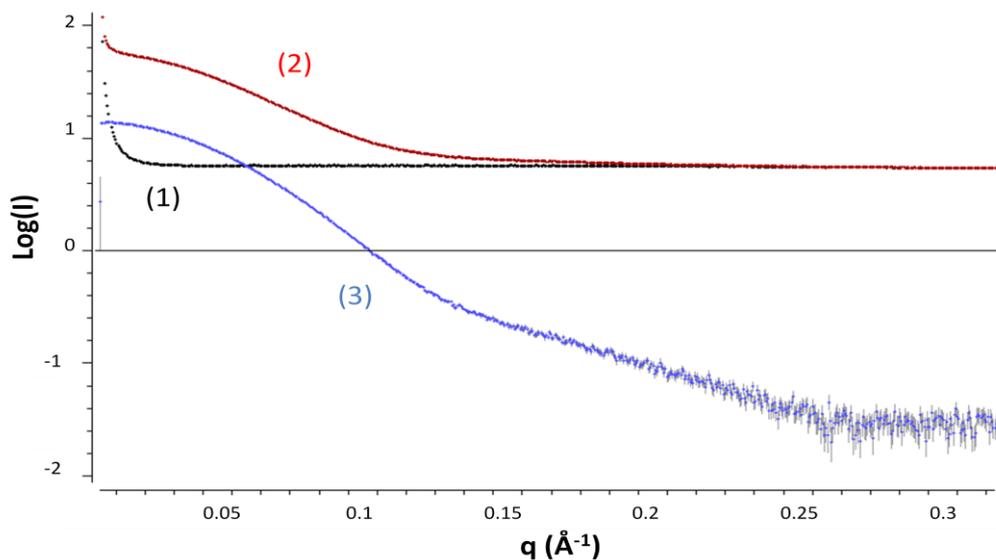
Where  $2\theta$  is half the scattering angle and  $\lambda$  is the wavelength of the incident beam. (In some literature and programs  $q$  can be referred to as  $s$  or  $h$ , sometimes with slightly different definitions (Putnam et al., 2007)). As the wavelength is kept constant, typically around  $\sim 1 \text{ \AA}$  and the scattering angle is confined within a low range, typically  $0.1\text{-}10^\circ$ , the  $q$  range measured in a SAXS experiment is approximately  $0\text{-}0.5 \text{ \AA}^{-1}$  (Koch et al., 2003).

The scattering function for an ideal solution of molecules is defined as:

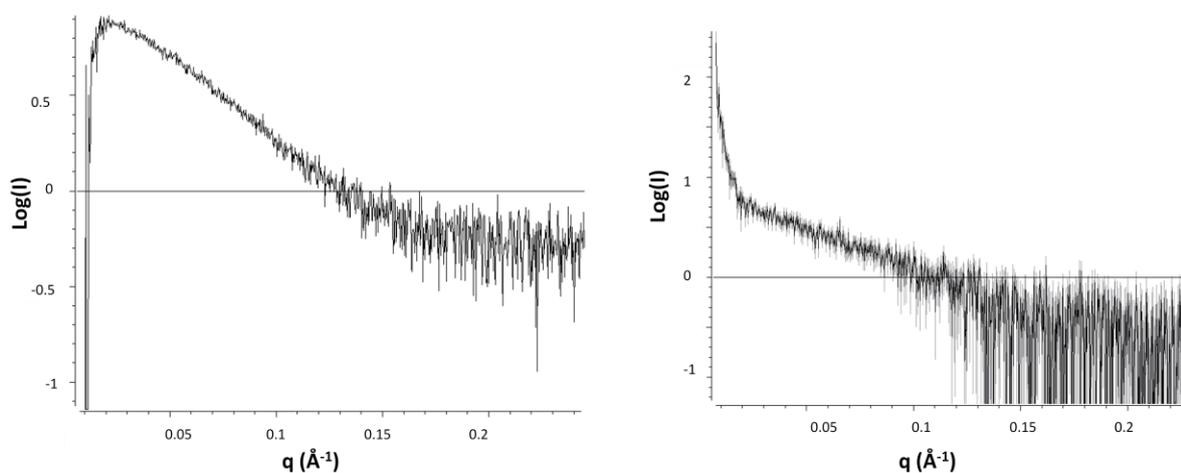
$$I(q)_\Omega = \langle I(q) \rangle = \langle A(q)A^*(q) \rangle_\Omega \quad (\text{Equation 5.2})$$

$A(q)_\Omega$  is a Fourier transformation over the excess scattering length density, and the scattering is averaged over all orientations, this average causes the  $I(q)$  function to be symmetric (Koch et al., 2003).

To obtain the SAXS scattering curve of a molecule it is necessary to subtract the scattering contribution of the relevant buffer in which the molecule is dissolved. The scattering contribution from solution is very large and it is thus very important that the buffer sample, used to measure its scattering curve, is identical to the buffer containing the sample (Koch et al., 2003). The sample container is also responsible for some contribution to total scattering. However, this can often be neglected since the same container is usually utilized for both sample and buffer and thus subtracted along with buffer contribution (Koch et al., 2003).



**Figure 5.2: Measuring SAXS scattering curves.** (1) and (2) shows the initial scattering curves for a BSA sample in buffer and the scattering curve of the buffer alone. (3) Displays the resulting scattering curve obtained by subtracting (1) and (2). It is apparent that the level of noise increases as  $q$  and thereby the angle is increased  
Figure produced in primusqt.



**Figure 5.3: Attractive and repulsive interactions.** Right: Sample in which  $SF(c,s) < 1$ . Attractive intermolecular interactions cause an increase in the scattering curve at low  $q$  range. These interactions may lead to unspecific aggregation which causes the sample to be inappropriate for further analysis. Left: Sample in which  $SF(c,s) > 1$ . Repulsive intermolecular interactions cause a decline in the scattering curve at low at low  $q$  range. Figures produced in Primusqt from insulin data.

Even a sample which is completely monodisperse may suffer from various effects of intermolecular interactions which will affect the scattering intensity at low angles. The scattering function for such a non-ideal solutions is given by:

$$I_{exp}(c, s) = I(q)SF(c, s) \text{ (Equation 5.3)}$$

SF(c,s) is the SAXS structure factor. Although similar in name, this is not to be confused the structure factor used in x-ray crystallography. The SAXS structure factor takes intermolecular interactions, repulsive or attractive into account. At infinite dilution at which no intermolecular interactions takes place, SF(0,q) = 1 (Koch et al., 2003). Repulsive interactions result in a structure factor value below 1 which cause the scattering to decrease (Figure 5.3, left), attractive interactions has the opposite effect (Figure 5.3, right).. The effects of minor repulsive and attractive interactions can be minimised by cropping data the data in the very low q range, attractive interactions may however cause unspecific aggregation which can invalidate the dataset (Koch et al., 2003; Putnam et al., 2007).

## 5.2 R<sub>g</sub> and I(0)

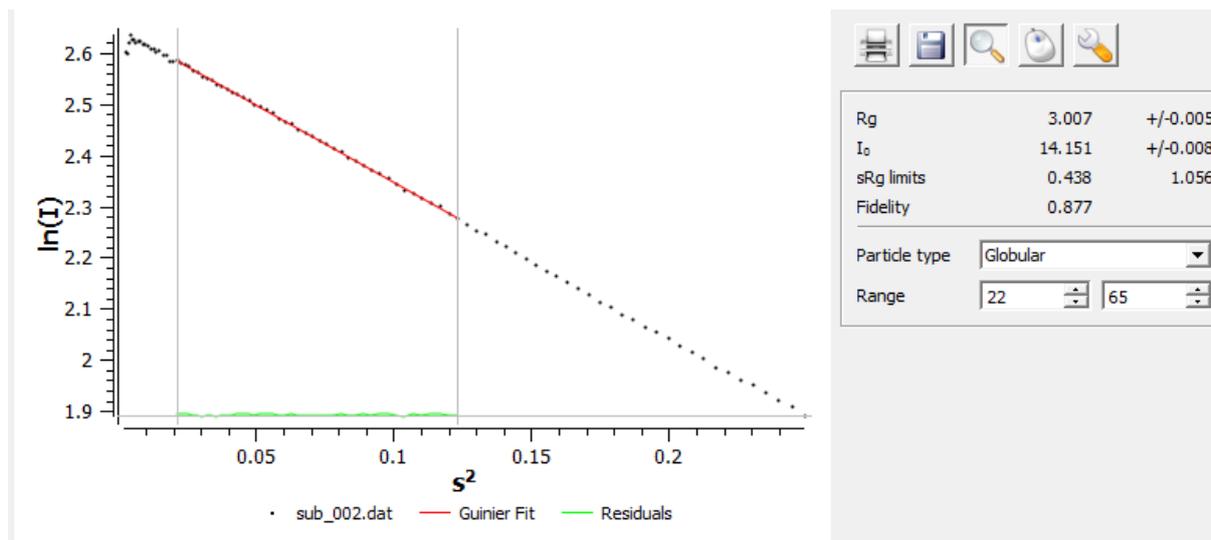
The most important parameters in a SAXS experiment are the I(0) and Radius of Gyration (R<sub>g</sub>) value. The R<sub>g</sub> value is the square root of the averaged squared distances of each scatterer from the particle center (Putnam et al., 2007). Large molecules will thus be characterized by large R<sub>g</sub> values. A sphere with a uniform particle density and a radius of 1 has an R<sub>g</sub> value of  $\sqrt{3/5}$  (Putnam et al., 2007), thus slightly lower (0.77) than its actual radius. The I(0) value is proportional to the number of electrons in the scatterer (Koch et al., 2003). If the exact concentration of the particle is known, I(0) can be used to determine the molecular weight. Both parameters are extracted from the Guinier approximation which at low q ranges can be given as:

$$I(q) = I(0)e^{-\frac{R_g^2 q^2}{3}} \text{ (Equation 5.4)}$$

This equation can be converted into a linear equation:

$$\ln(I(q)) = \ln(I(0)) - \frac{R_g^2}{3} q^2 \text{ (Equation 5.5)}$$

A plot of ln(I(q)) versus  $q^2$  will, at low q values in which the Guinier approximation is valid, result in a straight line (Figure 5.4). The slope of the curve yields the R<sub>g</sub> value while the I(0) is the intercept value. Since I(0) is the theoretical intensity measured at zero angle  $q=0$ , this value must be obtained from extrapolation of the curve (Putnam et al., 2007).



**Figure 5.4: AUTOGNOM Guinier region estimation in primusqt.** The experimental data points are shown in black, the Guinier fit in red and residuals in green.

### 5.2.1 Considerations

The Guinier plot is critical to SAXS analysis. While it is possible to obtain  $R_g$  and  $I(0)$  values by other means, lack of linearity in the Guinier plot usually indicates problem with the sample (Koch et al., 2003).

The Guinier approximation as only valid in the Guinier region, the region in which a plot of  $\ln(I(s))$  versus  $s^2$ , results in a straight line. This will not be the case at high  $q$  values, but any structure factors at low  $q$  values will also result in a non straight line.

The program Primus from the Atsas package contains an AutoR<sub>g</sub> function (Figure 5.4) which attempts to fit the Guinier region providing values for  $R_g$   $D_{max}$  and  $I(0)$ . While the parameters obtained serve as a good guideline, the program is, at the time of this, not completely reliable and thus all Guinier regions were determined and fitted manually.

Since  $R_g$  is simply the root of squared distances, impure samples may also display linear Guinier regions, and thus the Guinier plot is not suited to determine sample purity.

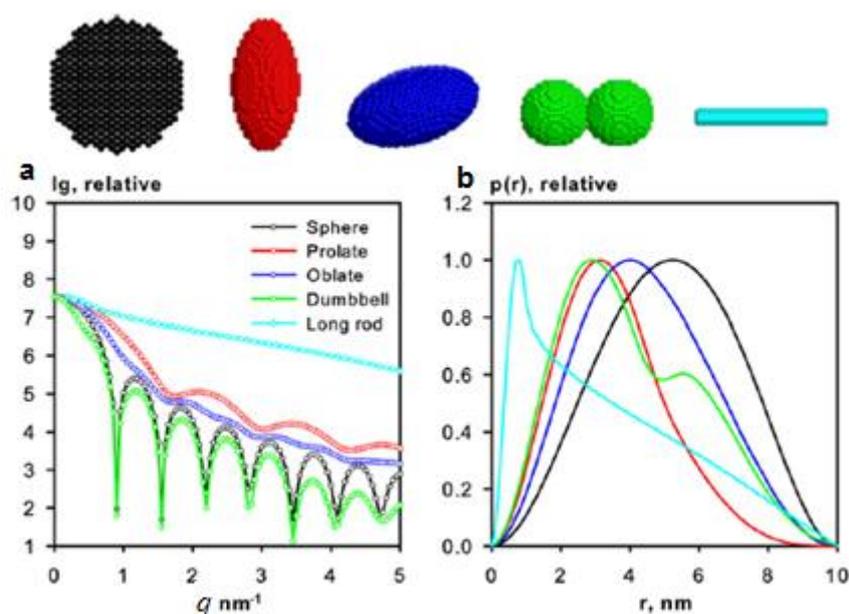
### 5.3 Pair distribution function (P(r))

Central to SAXS is the pair distribution function  $P(r)$ , similar to the Patterson function in X-ray crystallography (Putnam et al., 2007). The  $P(r)$  function is related to the scattering intensity as follows:

$$I(q) = 4\pi \int_0^{\infty} p(r) \frac{\sin(q * r)}{q * r} dr \quad (\text{Equation 5.6})$$

and represents the histogram of distances between pairs of scatterer within a particle (Koch et al., 2003). The  $P(r)$  function is used to determine the  $D_{\max}$  value, the maximum intra molecular distance (Koch et al., 2003). In theory the  $P(r)$  function will obtain a value of zero at  $r = 0$  and at  $r > D_{\max}$  (Koch et al., 2003). The calculation of a  $P(r)$  function is required for the 3d *ab initio* modelling of scattering data (Mertens and Svergun, 2010).

A  $P(r)$  function can be translated into simple overall shapes, some examples of these shapes and their corresponding  $P(r)$  functions are shown in Figure 5.5.



**Figure 5.5: Scattering intensities and pair distribution functions of common geometrical bodies with the same  $D_{\max}$  values.** (A): Scattering intensities. The curves describing the spherical and dumbbell conformations are ideal, in practise they will appear much smoother. (B):  $P(r)$  function. Elongated particles peak at an early  $r$  value corresponding to the cross section of the particle. Figure as published by (Putnam et al., 2007).

### 5.3.1 Considerations

The  $P(r)$  function requires input concerning which data points to include and what  $D_{\max}$  value to use. The data range used usually begins with the first data point used to determine the Guinier region to filter out data corrupted by interparticle interactions, and ended when noise levels are too high. The  $D_{\max}$  value is determined manually on a simple trial and error basis. There is still no universal accepted approach for the creation of  $P(r)$  functions. For this reason the obtained  $D_{\max}$  value should be interpreted with care.

## 5.4 *Ab initio* modeling

The purpose of *ab initio* modelling is to convert a one dimensional SAXS scattering curve into a three dimensional solution model.

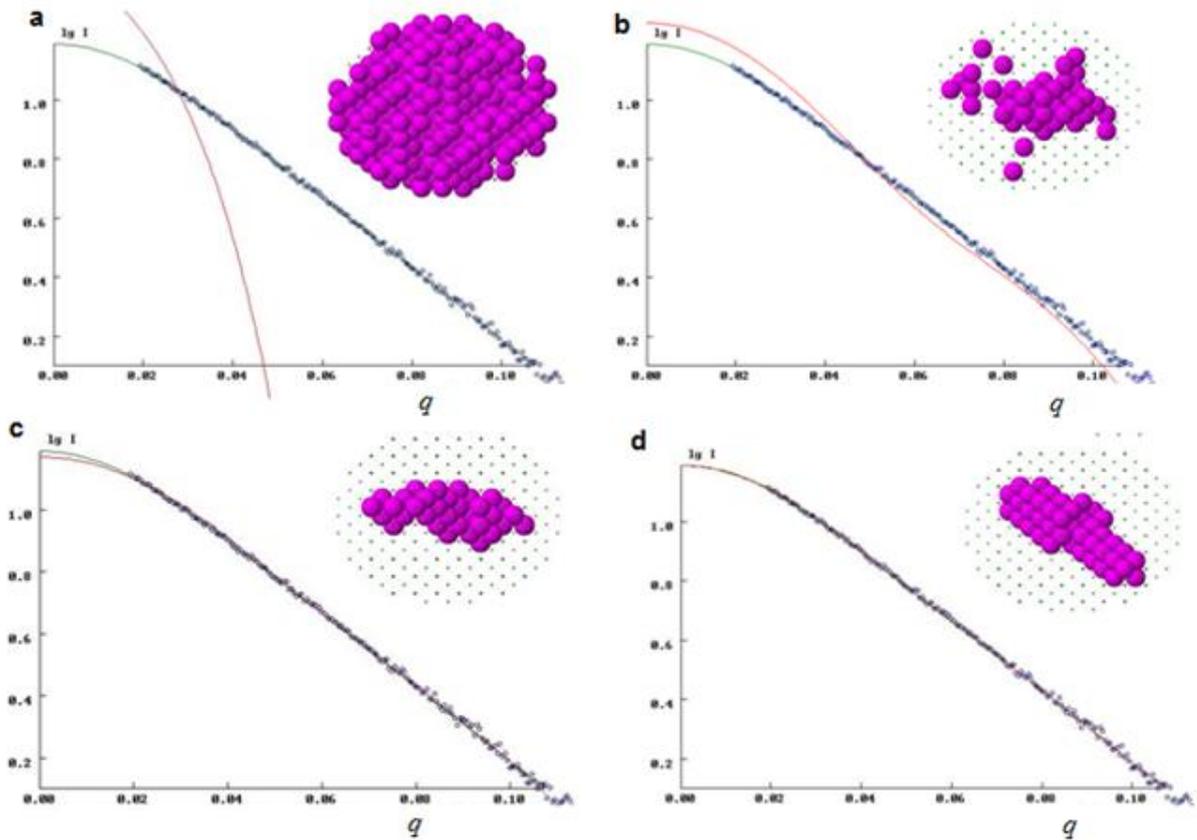
Initially a sphere with diameter equal or slightly larger than  $D_{\max}$  is filled with smaller beads (Mertens and Svergun, 2010). Each of these beads may either belong to the particle or to the solvent; these two states are indexed 1 and 0 (Koch et al., 2003). Starting from a distribution in which the entire sphere is packed, the model is modified to find the model that best fits the experimental data (Koch et al., 2003). Usually programs employ different types of Monte Carlo algorithms for this model fit (Koch et al., 2003). In a multicomponent system that contains components with distinctly different scattering lengths, such as a DNA binding proteins, it is possible to assign even further index values (solvent = 0 Protein = 1 DNA = 2) allowing the ability to distinguish the shape of the different components (Koch et al., 2003).

In practice *ab initio* model fit is achieved through simulated annealing (Mertens and Svergun, 2010). Random modifications are made to the system by changing index values, these changes will either move the theoretical scattering closer or further away from the experimental curve (Koch et al., 2003). Initially changes either way will be accepted but as the system is “cooled” the chance of accepting the second type of result will decrease (Koch et al., 2003).

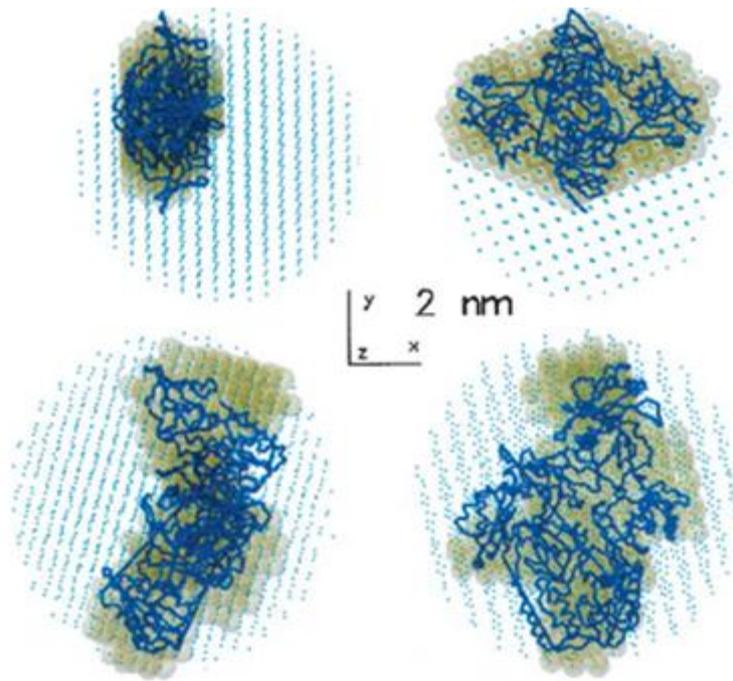
The *ab initio* models created in this thesis work were produced by the program DAMMIF, from the ATSAS program suite for *ab initio* modelling. The approach for this program is as follows(Svergun, 1999).

- 1) Start from a spherical configuration in which  $d = D_{\max}$  and at a high temperature  $T_0$
- 2) Select a bead at random, change its index value (from 0 to 1 or vice versa) and compute  $\Delta = f(\text{config}_{\text{new}}) - f(\text{config}_{\text{old}})$
- 3) If  $\Delta < 0$  always accept  $\text{config}_{\text{new}}$ ; if  $\Delta > 0$  accept  $\text{config}_{\text{new}}$  with probability  $\exp(-\Delta/T)$  otherwise accept  $\text{config}_{\text{old}}$ . Repeat step 2 with the accepted configuration.
- 4) Keep  $T$  constant for 100 reconfigurations or 10 successful reconfigurations whichever comes first, and then cool the system ( $T_{\text{new}} = 0.9 * T$ ) and repeat from step 1. Continue cooling until no further improvement in  $f(\text{config})$  is observed.

In order to decrease the number of possible structures, various constraints are imposed on the program such as penalties for unconnected or unpacked beads (Svergun, 1999). The bead model procedure is shown in Figure 5.7.



**Figure 5.6: *ab initio* modelling in the program DAMMIN.** Modelling of a helical oligomer in DAMMIN. (a-d) represents 1, 10, 50 and 120 iterations respectively. The beads accepted in the structure are shown in pink, their resulting scattering curve is shown in red. The experimental scattering curve is shown in green. The model was accepted after 120 iterations. To increase speed a packing radius of 10 was used resulting, this results in much fewer beads than normally used. Only a small portion of the curve is fitted. DAMMIN is similar to DAMMIF but contains a graphical interface and was used for illustrative purposes. The Figure was produced using elongated insulin oligomers.



**Figure 5.7: Examples of *ab initio* generated models:** Models and crystal structures of thioredoxin (left) and reverse transcriptase (right). The light blue dots indicate the initial search volume. The yellow dots are the model determined by DAMMIN. The blue lines are the structures determined by X-ray crystallography. Figure adapted from (Svergun 1999).

### 5.4.1 Considerations

Using *a priori* knowledge about the structure, various constraints, such as symmetry, can be introduced into the model. While this can generate models which appear to be of higher quality they may not reflect the real structure. Figure 5.7 depicts some solution structures obtained from SAXS experiments along with the relevant crystal structures superposed. These envelopes are of high quality, and are the best one can expect to be able to produce, without the introduction of constraints.

In the modelling procedure it is possible to use beads of different sizes. Smaller beads increase the resolution of the resulting envelope but increases the required simulation time. Larger beads have the opposite effect. While it may be tempting to use very small bead sizes, the final model resolution is usually limited by the quality of the scattering curve.

DAMMIF will attempt to model any  $P(r)$  function and does not have the functionality to recognize a polydisperse sample.

Usually twenty envelopes are determined and averaged. If envelopes show too large a discrepancy from the average envelope they are discarded. *Ab initio* envelopes should always be critically evaluated, especially if samples are suspected of being polydisperse. Even small amounts of aggregated proteins will cause the generation of an invalid model.

## 5.5 CRY SOL fitting of experimental and theoretical scattering curves

CRY SOL enables one to derive a theoretical scattering curve from a known crystal structure, to compare this to an experimental scattering curve and to calculate a deviation in form of the  $\chi^2$  value (Svergun et al., 1995). CRY SOL thus allows one to test if a scattering curve matches a known crystal structure. Due to the flexibility of a protein in solution, the experimental scattering curve will rarely completely match the derived scattering curve.

The  $\chi^2$  is calculated as:

$$\chi^2 = \frac{1}{N-1} \sum_j \left[ \frac{I_{exp}(q_j) - cI(q_j)}{\sigma(q_j)} \right]^2 \quad (\text{Equation 5.7})$$

In which N is the number of experimental points, c is a scaling factor and  $\sigma(s_j)$  is the experimental error at the momentum transfer  $s_j$  (Petoukhov and Svergun, 2005).

$\chi^2$  values should be interpreted with caution; if the errors in the measurements of the experimental data are high this will result in a low  $\chi^2$  value. Thus scattering curves of poor quality will appear to resemble known crystal structures much more frequently than high quality structures. This also impairs comparisons between different experimental curves in which the error values will be different. CRY SOL also adds a hydration shell to the crystal structure.

## 5.6 Rigid body modeling

If a complex consists of K subunits with known structure and their partial scattering amplitudes can be denoted  $A_{lm}^{(k)}(q)$  and the scattering intensity of the entire complex can be described as:

$$I(q) = 2\pi^2 \sum_{l=0}^{\infty} \sum_{m=1}^l \left| \sum_{k=1}^K A_{lm}^{(k)}(q) \right|^2 \quad (\text{Equation 5.8})$$

As is the case with *ab initio* modelling, simulated annealing is applied in rigid body modelling. While several programs can be used for this purpose, in this thesis work only the program SASREF was used. The program starts with a fixed initial assembly of the K subunits and computes the scattering intensity of the complex using the above equation. Each subunit is then rotated by an arbitrary angle and shifted along an arbitrary direction and a new scattering intensity is calculated. The rigid body fitting of known models to experimental SAXS curves, the following function is minimized.

$$E = \sum (X^2)_i + \alpha_{dis}P_{dis} + \alpha_{cross}P_{cross} + \alpha_{cont}P_{cont} \quad (\text{Equation 5.9})$$

Here  $\chi^2$  is calculated as in equation 5.7 and penalty values are added for disconnects ( $P_{dis}$ ), crossovers ( $P_{cross}$ ) and lack of continuity ( $P_{cont}$ ) (Petoukhov and Svergun, 2005).

SASREF adds hydration shells individually for each subunit whereas CRY SOL adds a hydration shell for the entire particle. This causes SASREF and CRY SOL fits to give different  $\chi^2$  values for the same identical structures and SASREF models should thus have their  $\chi^2$  value recalculated in CRY SOL for comparison purposes.

