UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

School of Chemistry

The SmartLab: Experimental and Environmental Control and Monitoring of the Chemistry Laboratory

by

Stephen M. Wilson

A thesis for the degree of Doctor of Philosophy

June 2011

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS
SCHOOL OF CHEMISTRY

Doctor of Philosophy

The SmartLab: Experimental and Environmental Control and Monitoring of the Chemistry Laboratory

by STEPHEN M. WILSON

The work presented in this thesis focuses on acquisition of data such as environmental and experimental conditions and the control of experimental apparatus. A suite of software applications will be described which captures data from a standard data acquisition card, processes this and archives it in an appropriate repository. The software revolves around a central message broker, controlling the flow of messages between intercommunicating components.

The automation section will focus on the control of many components of an experiment studying the air/liquid interface using the technique of second harmonic generation (SHG). Software will be described that can be used to controls actuators, laser sources and other hardware within the laboratory. This software will use the same message broker to send and receive messages relating to the devices they are controlling.

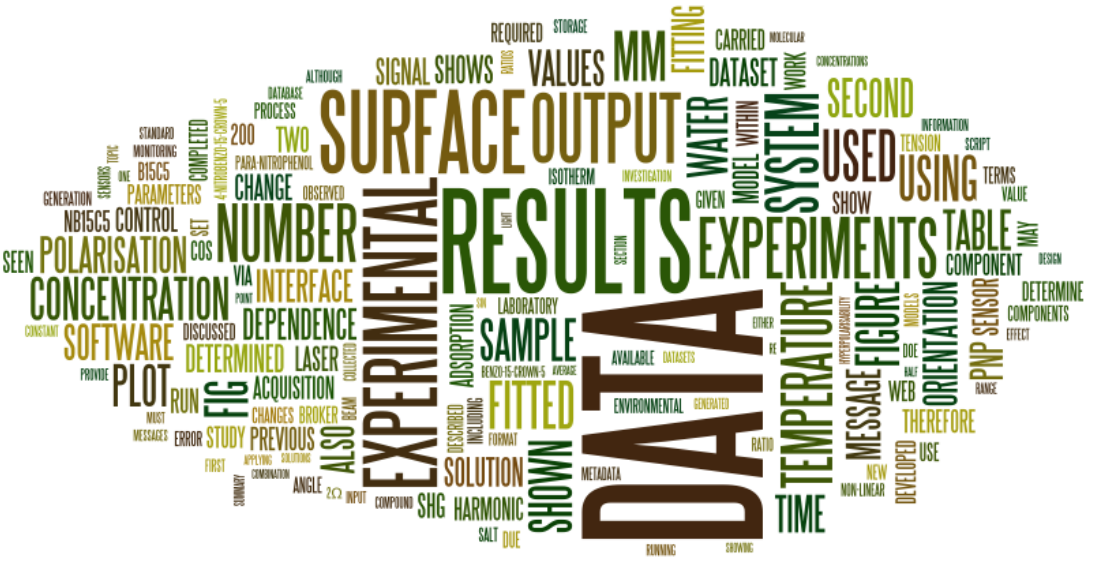
These tools were used to run a number of SHG experiments studying the air/liquid interface of para-nitrophenol (PNP), benzo-15-crown-5 (B15C5) and 4'-nitrobenzo-15-crown-5 (NB15C5). The complete experimental process was automated and results made available through a web interface, allowing real-time monitoring and decision making of each experiment.

Acknowledgements

I would like to thank Professor Jeremy Frey for his support and guidance throughout the work. I would also like to thank the EPRSC for funding the research. Dr Leftris Danos, Dr Andy Standord-Clark and Dr Robert Greef have also provided valuable support during my research.

In addition, thanks go to the Frey group for their support with any coding issues and the Stulz group for the use of their laboratory. Special thanks go to Tom for the use of his laboratory desk space.

Finally, many thanks go to my family and loving fiancée for their understanding and support during the numerous years of my student life.



Contents

List of Figures

List of Tables

Abbreviations

API application programming interface

B15C5 benzo-15-crown-5

CSV comma separated variables

DOE design of experiment

ELN Electronic laboratory notebooks

GMT Greenwich mean time

GUI graphical user interface

I/O input/output

IP intellectual property

LIMS Laboratory information management systems

MQTT Message queue transport telemetry

NB15C5 4'-nitrobenzo-15-crown-5

OGC Open Geospatial Consortium

OGSA Open Grid Services Architecture

PDF portable document format

PHP PHP: hypertext preprocessor

PMT photon multiplier tube

PNP para-nitrophenol

REST representational state transfer

SFG sum frequency generation

SHG second harmonic generation

SOS sensor observation service

SWE sensor web enablement

TCP/IP transmission control protocol/internet protocol

VAS variable angle spectroscopic

VI virtual instruments

VNC virtual network computing

WSDL Web Services Description Language

XML Extensible markup language

Chapter 1

Introduction

1.1 eScience

The overarching goal of this work was to investigate e-Science techniques and technologies, focusing on the chemistry domain. The application of these technologies will improve the experimental process, increase the volume and quality of reported data and allow distributed collaboration between research groups across the world.

There are several definitions of the term e-Science:

> "The term e-Science denotes the systematic development of research methods that exploit advanced computational thinking"[1]

> "eScience refers to the large-scale science that will increasingly be carried out through distributed global collaborations enabled by the Internet"[2]

> "E-Science is the tool that offers scientists a scope to store, interpret, analyse and network their data to other work groups"[3]

This is best highlighted in a diagram produced as part of the oreChem project[4], shown in Fig. 1.1. This diagram shows how many sources of data can be collected by a number of independent software applications. Using a common communications protocol these individual components become interoperable, allowing data to be shared and re-used across a wide range of applications and users.

These definitions of eScience highlight the need to develop computational tools to support the entire research life cycle; data acquisition, storage, analysis and publication. A number of areas of research have branched from the eScience domain looking to support different aspects of this research life cycle. The tools to support the eScience goals are

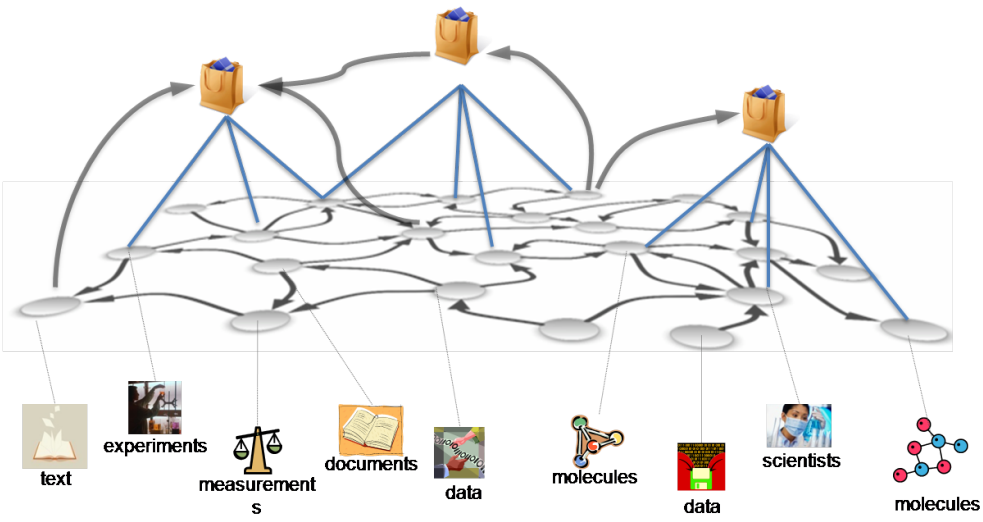


FIGURE 1.1: The oreChem project diagram, showing many integrated components communicating together. Src: http://research.microsoft.com/en-us/projects/orechem/

now becoming available due to increases in technology, such as fast networking and improved CPUs. The development of web service standards, based on XML, has allowed many different system to integrate together through machine-to-machine interactions. With the development of Web 2.0 technologies, the traditional static web pages have been replaced by interactive interfaces to the data. Dynamic content can be accessed, in real-time, from data stores and displayed in a wide variety of formats. With the introduction of Smartphones these web based solutions can be access on hand-held devices, allowing real-time access to the laboratory data from any location[5].

1.1.1 Laboratory information management systems

Laboratory information management systems (LIMS) support the process of data acquisition from an experiment, either simulation or physical, through to data storage and archiving. The system should capture not just the experimental results but the metadata associated with that experiment, such as the samples, instrumental parameters, the environment conditions, experiment processes, and the researcher enacting those processes. Through use of web-based technologies, this data can be made available from any location, in real-time, and from any operating system. LIMS are typically used in large-scale analytical laboratories where extended volumes of data are collected. Many analytical techniques may be carried out on a single sample. By automatically linking experimental results to a sample all output of the given sample are available from a single location.

Where large volumes of raw data is collected, beyond what the researcher is capable of handling, the LIMS software can provide the first level of work up and analysis, reducing the data down to a manageable size. These tools should be developed in such a way

that the raw data, or at least the analytical process, is available so the researcher can revert back to the original data if required.

A typical chemist will have information stored in a number of distinct computers spread throughout the department; NMR spectra and mass spectrometry results on dedicated machines and analysed data on a personal computer. This can lead to data being lost or, if poorly recorded, misinterpreted. The LIMS software should also provided a central point of storage for all laboratory data. This does not just provide a structured solution but can also link data automatically if the correct metadata is collected.

There are a number of commercially available LIMS packages. SQL*LIMS^{TM}[6] provides an enterprise level suite of web-based software aimed at a wide range of industries. The software is designed to manage the entire process life cycle (tracking samples, data analysis and results sharing) while integrating with existing software. BioRails^{TM}[7] is a LIMS product aimed at the biological disciplines. This uses Web 2.0 technologies to develop a modular system for archiving data (including IP protection), chemical indexing, work flow development and scheduling. The software separates the plan and realisation of a work flow, allowing deviation between the two as occurs during scientific research. BioRails^{TM}also provide a number of supplementary software packages for linking into other software such as Microsoft Office^{TM}[8]. Labtronics^{TM}[9] have developed LimsLink^{TM}, a software solution for integrating laboratory instruments to a LIMS interface. The software offers automated real-time data transfer between instruments and the LIMS interface, as well as linking to other software solutions used in the laboratory. There are many others including LabVantage^{TM}(http://www.labvantage.com/), LabCollector^{TM}(http://www.labcollector.com/), BTLIMS^{TM}(http://www.btlims.com/), Nautilus^{TM}(http://www.nautilus.com) and LabSoft^{TM}(http://www.labsoftlims.com/).

Many of the existing LIMS packages are both expensive and developed for a specific task or research domain. This makes then difficult to be tuned for a new domain. More generic solutions do exist but can be more time consuming to configure and maintain.

1.1.2 Electronic laboratory notebooks

Electronic laboratory notebooks (ELN) aim to improve the way experimental practises and results are recorded. The fundamental aim of an ELN system is to provide a reliable, reproducible and standardised replacement to the traditional paper notebook. It has been suggested that in the chemistry domain, laboratory notebooks often do not contain all the required information to repeat an experiment[10]. Typically a notebook will contain an experimental plan, with associated safety information, written before the experiment is run. During the realisation of that plan, key values such as weights of sample and colour changes are recorded. Once an experiment has completed printed copies of analytical instrument output, such as mass spectrometer results, are stuck into

the paper book. Where print-outs are not available the results may need to be copied in by hand.

Through use of an ELN, this process can be improved greatly. The ELN software must therefore include the traditional functions of the notebook, storing the plan and thoughts of the researcher, while capturing and linking data from a multitude of sources[11]. The provenance of the data must also be considered when developing an ELN; in recording the ownership of the data, the companies intellectual property (IP) rights are protected. This can achieved using techniques such as digital signatures. A key feature which extends the ELN from a paper notebook is the ability to search records. Should the experimental metadata have been recorded and linked correctly during the experiment, more advanced searches can be carried out and produce more relevant search results.

The major consideration when developing an ELN system is data archiving. In the pharmaceutical industry a drug can take up to ten years from concept to reaching the market[12] and in health and safety, data records must be made available for up to fifty years[13]. Over this time-scale, standard office software can undergo a number of revisions that may not be backwards compatible; when developing an ELN this must be accounted for. The portable document format (PDF) file format can overcome this problem as its representation is independent of the hardware and software. The use of XML can also help with future proofing, as well as aiding interoperability via the use of an application programming interface (API)[14].

There have been major advances in hardware design that has aided the uptake in ELN software. Devices such as tablet PCs and Smartphones allow users to take touch and pen sensitive devices into the laboratory and record experimental procedures as they are carried out. As many ELN systems use a centralised database to store the data, should the device be damaged no data is lost. In contrast, should a traditional paper notebook be damaged, such as through a solvent spillage, a considerable volume of data may be lost.

As with the LIMS software, there are a number of commercial ELN products available for purchase. Labtronics^{TM}have developed Nexxis ELN^{TM}[15], part of the same suite of software as LimsLink^{TM}. This system is aimed at highly structured experiments which are repeated regularly. An experiment is presented as a web form, with boxes to fill in the relevant experimental details. The process of completing this form can be automated, through retrieval from experimental apparatus and the data repository. The use of structured forms promotes consistency in the way experiments are recorded, making sharing and collaboration easier between users. The software is web-based, and therefore independent of the client running the software (although at present limited to Internet Explorer). As the ELN is developed as part of a suite of software, it can easily be integrated with the other components in the package. This enhances the usability of each component as data sharing and linking can be automated, although if

developed correctly this should be possible between independently developed software. IDBS have produced E-WorkBook™[16], an ELN developed for any science domain. The software offers a framework for developing specific solutions while promoting data sharing and re-use. E-WorkBook also provides a high level of IP protection. This is achieved through use of SAFE signatures[17], an independent digital authentication association, and capture of document audit trails. This audit trail includes all entries, modifications and deletions with the ELN. As well as the generic ELN, IDBS™also provide a chemistry extension; offering functionality specific to the chemistry domain. E-Notebook™[18], from CambridgeSoft™, follows a different approach to acquiring the scientific date. Data is collected from existing document types such as Microsoft Office files, PDFs and images. Sharing and collaboration is supported through shared drives and extensive search routines. As E-Notebook is part of a larger suite of software, including ChemDraw, chemical data can be automatically generated and imported into the system with the associated metadata.

LabTrove™[19] is a ELN in development at the University of Southampton. The research notebook is represented as a blog, a new post representing a new entry to the notebook. The unstructured nature of a blog allows storage of a wide range of data types and information. Where structure is required, templates can be used to dictate the format of the post. The blog solution promotes collaboration through commenting, allowing other users to discuss and share ideas around a given post. LabTrove extends the standard blog implementation as it contains a data repository. Data files can be associated with blog posts and embedded within the post itself. The ELN handles IP protection by recording a posts audit trail, storing all modifications.

As with the LIMS software there are many more ELN systems that have not been discussed, including Amphora™(http://www.amphora-research.com/), Symyx™ (http://www.symyx.com/), LabTracks™(http://www.locustechnology.com/labtracks) and Contur™(http://contur.com/).

1.2 Interfacial Chemistry

An interface is defined as the boundary between two bulk mediums. They occur everywhere, formed at the surface between air, liquids, solids and vapours. It is important to understand the interface between two mediums as molecules present at the interface have unique chemical and physical properties different to that of the bulk[20, 21]. These can include pH, transport, molecular geometry and dielectric properties[22]. Any transfer between the two mediums, such as drug delivery or ion transport, must pass through the interface and therefore an understanding of these properties is vital. The generation of the interface has an associated free energy of formation, which is a measure of the energy required for a molecule to be at the interface.

The importance of interfaces can be highlighted by looking at the range of scales they occur at. This can be as small as the cell wall, where enzymatic reactions control the transport of molecules in and out of the cell. On a much large scale, interfaces also have importance in the ozone where gasses are broken down on the surface of ice crystals. There are many other important interfacial processes, such as corrosion, heterogeneous catalysis, ion-exchange and electrochemical processes, that have significant consequence on the world we live in. In understanding these systems we can design and control these processes.

Study of interfaces can be difficult using traditional spectroscopic techniques due to the large signal from the bulk. It is therefore vital to use methods selective to the surface. Techniques such as scanning tunnelling microscopy (STM) and Auger spectroscopy are used to determine surfaces of solids but are not suitable for liquids. There are a number of non-linear optical techniques which can be used to study a liquid surface, including SHG and sum frequency generation (SFG). These methods are advantageous as they are both selective to the surface and non-destructive; allowing *in situ* study of the interface. The use of these techniques allows properties such as molecular orientation, surface free energy and surface tension to be determined. A number of techniques used to study the liquid interface are outlined below.

1.3 Techniques for study of the liquid surface chemistry

1.3.1 Surface Tension

In a liquid bulk, each molecule has equal forces acting upon it in all directions. At the interface this symmetry is broken, resulting in molecules at the surface having a higher potential energy. The liquid will re-arrange to minimise this energy by minimising the surface area and maximising the volume of molecules in the bulk.

Surface tension, γ, is defined as *the measure of force acting at right-angles to any line of unit length on the surface*[23] and is typically reported in units of Nm^{-1}. The surface tension of a pure liquid is typically well defined, as such any variation from this known value can indicate the presence of impurities at the surface. As this is a macroscopic measurement, it cannot identify specific compounds at the surface. There are a number of techniques used to measure surface tension, including the Wilhelmy plate, bubble pressure, capillary rise and Du Nouy methods.

1.3.2 Second Harmonic Generation and Sum Frequency Generation

Non-linear optical techniques such as SHG and SFG can be used to study the liquid interface[24, 25]. Both SHG and SFG are electrically forbidden in centrosymmetric

media such as bulk liquids. At the interface this inversion symmetry is broken, therefore signal from both techniques can be observed. This phenomenon makes these techniques highly surface selective, even allowing observations of the surface properties at sub-monolayer coverages.

In SFG two input photons of frequencies ω_1 and ω_2 and converted to a signal output photon of frequency ω_3, where $\omega_3 = \omega_1 + \omega_2$. SHG can be considered a special case of SFG where $\omega_1 = \omega_2$, therefore requiring only a single laser source. SHG has been the primary method for investigating liquid/air interfaces during with work and will be discussed further in chapter 6.

1.3.3 Ellipsometry

Ellipsometry is an experimental technique used to study thin films at a sample surface[26]. A beam of light at a known polarisation is reflected from a sample. Changes to the polarisation and amplitude of the light on reflection are monitored. From these measurements two parameters can be determined, delta and psi. Delta is the phase difference between the S and P components of the input beam and the S and P components of the reflected beam, ranging from 0^o to 360^o . Psi is the change in amplitude between the input and reflected beam. These parameters can be used to determine surface properties.

Spectroscopic ellipsometry can be used to determine how the complex refractive indices changes with wavelength. This is carried out by scanning through a range of wavelengths of the input beam and observing the changes in delta and psi with respect to wavelength. The morphology of the surface can also be observed using imaging ellipsometers, an extension of a standard ellipsometer where the detector has been replaced by a CCD camera.

Chapter 2

Environmental and Experimental monitoring

2.1 Background

Experimental conditions have always been reported as part of the experimental procedures. Including this information allows the experiment to be repeated by researchers who have little background on this type of experimental chemistry. Even if followed exactly, the execution of these procedures may not always produce the expected results. It has been reported that this is often due to incomplete procedures being reported due to poor laboratory notebook keeping[10]. Although this may be true in some circumstances this is not always the case. The notebook may contain all factors measured as part of the experiment but not include environmental factors that may have affected the experimental procedure, such as humidity, light levels or air flow. The most prevalent of these factors is room temperature. This can be affected by a number of factors including air conditioners, global location or other experiments carried out in the local vicinity. As well as providing more metadata to reported procedures this environmental data may also highlight why an experiment which has worked in the past, failed on a given run.

Although it is extremely useful to have this data available, it is difficult to ensure a chemist will record it all manually. The use of an automated data capture system through pervasive computing can provide a solution. The system should run independent of an experiment, allowing comparison of conditions during and outside the experimental time-frame. A system such as this can also be used to trigger alarms should conditions change outside the expected range, which can aid in the safety of the chemists and also catch potential failures of experiments earlier than if the monitoring was not in place. To increase the confidence in this data the entire data life cycle should be captured as part of the acquisition system[27]. This should include the metadata on the sensor, such as model, manufacturer, calibration methods, sensitivities, etc...

As discussed later in later chapters, this project has focused on a specific experiment; the study of the air/liquid interface via second harmonic generation. The results of this experiment are particularly sensitive to changes in temperature as well as movement in the laboratory causing ripples on the liquid surface. There are also humidity factors to consider which will affect the rate of evaporation of the sample.

A number of systems have already been developed to varying degrees, each solutions advantages and disadvantages will be discussed. Sanchez *et al.* describe a data acquisition system for capturing data from the TJ-II tokamak[28]. In this example an existing system is in place using custom C/C++ software to capture data and store it into an SQL database via Berkeley sockets. A stand-alone graphical user interface (GUI) is offered to configure the acquisition and provide real-time monitoring. It discusses how strong coupling between point of acquisition and the repository makes it difficult to add new channels (both acquisition and storage software modifications). A new system is proposed using LabView[29] and a number of National Instrument data acquisition cards to provide a low cost modular data acquisition system. LabView is also used to generate new GUIs which are controlled via virtual network computing (VNC) software[30]. The use of modular components allows modifications in the acquisition system to be made much quicker, with only a small sections of the system require modifications.

This problem of strong coupling between acquisition device and storage server has been widely discussed[25, 31, 32]. Each suggest the use of middleware to provide a dynamic link between data producers and the storage solution. The middleware decouples the components of the system, each element unaware of the whole system and only communicating with a single central point. This can run over a number of communication channels such as controller area network (CAN) serial bus[33] or transmission control protocol/internet protocol (TCP/IP)[34], among others.

A number of solutions make use of web-enabled systems to provide benefits such as reduced administration costs[35], easy access to data[36] and through additional technologies provide control over equipment[37]. Klimchynski *et al.* discuss a data acquisition system which generates dynamic web pages via an embedded web server[38]. The sensor data is presented in real-time via an internet browser using JavaScript. The use of the web technology provides access to the data from anywhere with a connection to the internet. As this system uses a micro-controller to generate the web pages, it is unable to store and therefore recall previous data. To overcome this a centralised repository could be used to archive the data, a separate web server can recall and present the data via dynamic web pages. Through the use of an API external software can also remotely access and retrieve the data directly; this is discussed further in chapter 3.

2.2 Overview of proposed system

A system is proposed to capture, process and store data from any standard sensor deployment, focusing on experimental and environmental data within the chemistry laboratory. The system will revolve around a central piece of software supplied by IBM[39] called the Microbroker. The Microbroker is a message broker controlling the flow of data using the Message queue transport telemetry (MQTT)[40] protocol and uses the publish/subscribe architecture to provide the link between the source of the data and its destination. Standard sensory equipment such as off-the-shelf temperature, humidity and movement sensors will be connected to a data acquisition card. The data captured by the data acquisition card will then be published to the Microbroker, undergoing the data life-cycle before it is stored in a central repository or received by an output device. Where possible experimental set-ups will also be connected to the system through standard data communication channels such as RS232 or through a manufactured dongle, capturing the data these systems output and publishing it through the same standardised communication channels.

This data will be made available to the users through a number of systems such as an API for machine-to-machine communications, a website for on-line data recall and as real-time streams on a number of client computers and output devices. An alarm system will also be implemented to provide alert emails to users should the data captured moved outside of expected values (defined by the users).

Additional to this, separate systems will be implemented to capture results from devices that do not provide data during the experimental process. In this situation an application will be produced to monitor when files have been written to the computer and automatically upload the file to the central repository. As these files may be large (compared to the size of a standard MQTT message) they will be uploaded to the central repository via a separate channel, saving bandwidth and processing power of the server running the message broker.

2.2.1 Publish/Subscribe technology

The publish/subscribe model dissociates the components that produce the data from those who consume it[41]. This is achieved using a central broker which all clients connect to. Data producers will publish their data to the broker with an identifier, determining the ‘type’ of the data. The broker determines which consumers are subscribed to those ‘types’ and transfers the messages accordingly. Any client in the system may be both a publisher and subscriber.

The Microbroker uses a topic based filtering system to determine the message ‘type’. To publish a piece of data the application must first create a TCP/IP connection to

the MicroBroker. Data is published to the MicroBroker similarly to that of an email; the message is split into two parts, the content of the message containing the data and a topic that the message is related to. To receive a published message an application must again create the TCP/IP connection to the broker but must then subscribe to a topic. When a message is published, the Microbroker will determine which clients are subscribed to the message topic and forward the message to those clients. Messages can also be published with a retained flag; when these messages are sent the Microbroker will hold the most recent message published on that topic. If a client then subscribes to that topic the retained message is published once the subscription is made.

Topics are organised in a hierarchical system using '/' to donate new levels in the hierarchy. When a client subscribes it can either be to a single topic within the hierarchy or using 'wild-cards' to a range of topics. The wild-cards available are:

- # – The # wild-card allows you to subscribe to all levels of the hierarchy below specified level, shown in Fig. 2.1.

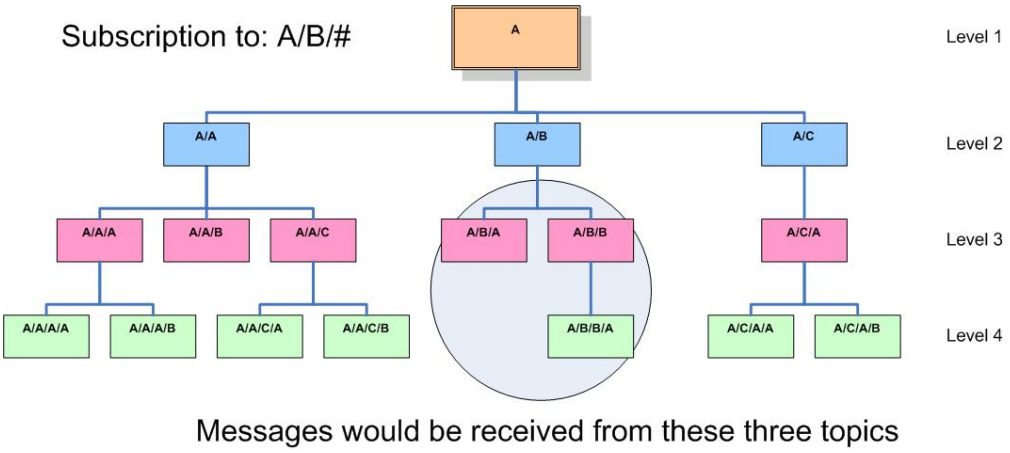


FIGURE 2.1: Subscription to the # wild-card

- + – The + wild-card allows you to skip a level of the hierarchy, subscribing to all lower levels tiers from a range of the given level, shown in Fig. 2.2

2.2.2 Message Queue Transport Telemetry

Message queue transport telemetry (MQTT) is an open source protocol used in combination with the IBM Websphere MQ software using the TCP/IP transport layer[40]. Its key design specifications are to be lightweight, allowing the use over low-bandwidth networks, and for it to be simple to implement onto embedded systems. The protocol has three levels of quality of service the publishing client may select before sending the message, determining how the Microbroker responds on arrival[42]:

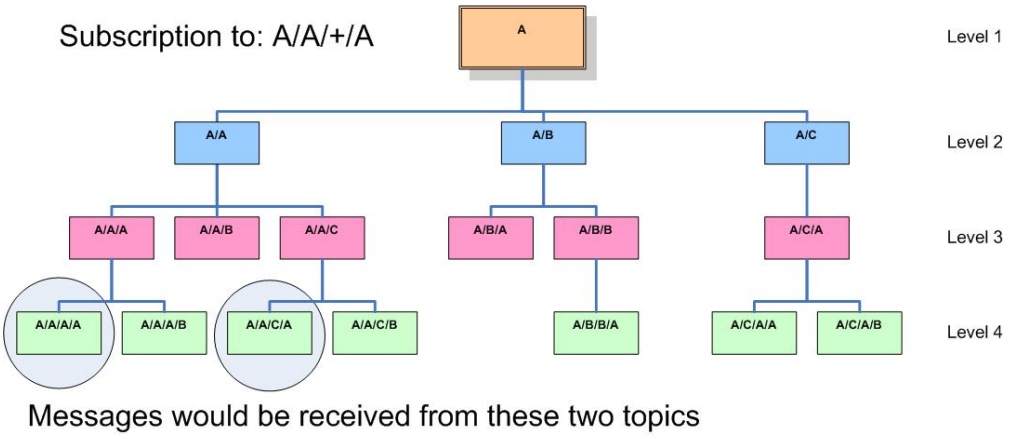


FIGURE 2.2: Subscription to the + wild-card

- Quality of Service 0 – Single publish without any checking that it arrived.
- Quality of Service 1 – Publish and await confirmation of delivery message. If the confirmation message is lost the message will be resent until the confirmation is received.
- Quality of Service 2 – Publish as part of a conversation between client and broker; this guarantees once and once only delivery.

The protocol also provides a service to monitor the connection between client and broker. When a client subscribes to the broker they can publish 'special' subscribe message with a keep alive-time so that if no messages are received within this time the broker assumes the connection has been lost and publishes a notification on an administration topic.

The message are kept lightweight by using a two byte header. The first byte is used to contain message information such as type and quality of service. The second byte contains information of the message length (including the header). In 'special' messages a variable header may also be present to determine additional parameters relating to the message[43].

2.2.3 Middleware

Middleware is the name given to software that connects two components while keeping them independent[44]. It provides a set of common procedures for communicating and transferring data between those components, while hiding the low-level programming. The use of middleware allows the developer to abstract themselves from the data transfer mechanisms, allowing more time producing functionality within the application.

In this system the middleware will provide the connection between the sensors and the repository or output devices. This abstraction between the components makes the

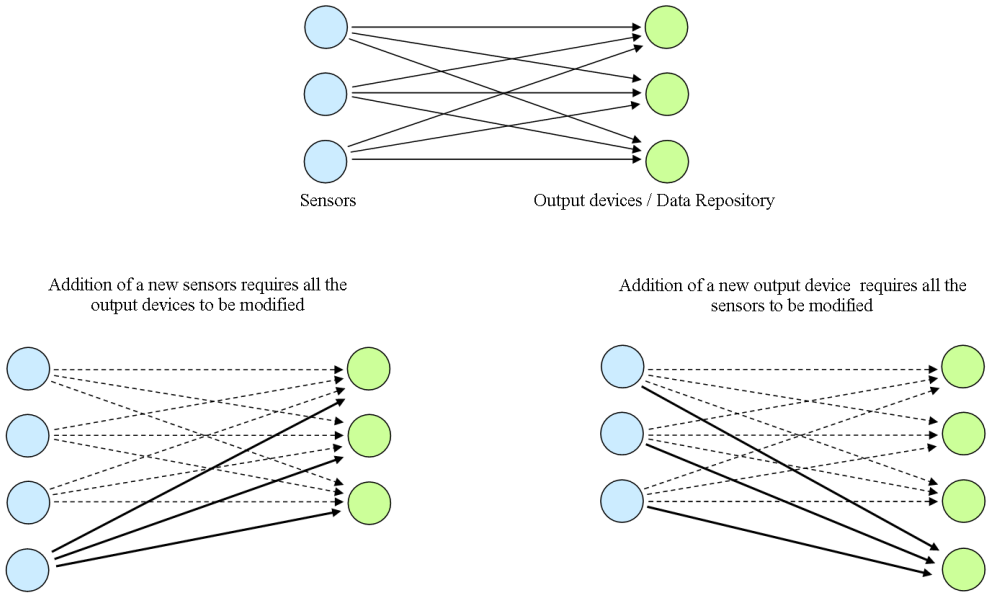


FIGURE 2.3: Addition of a new component without using middleware

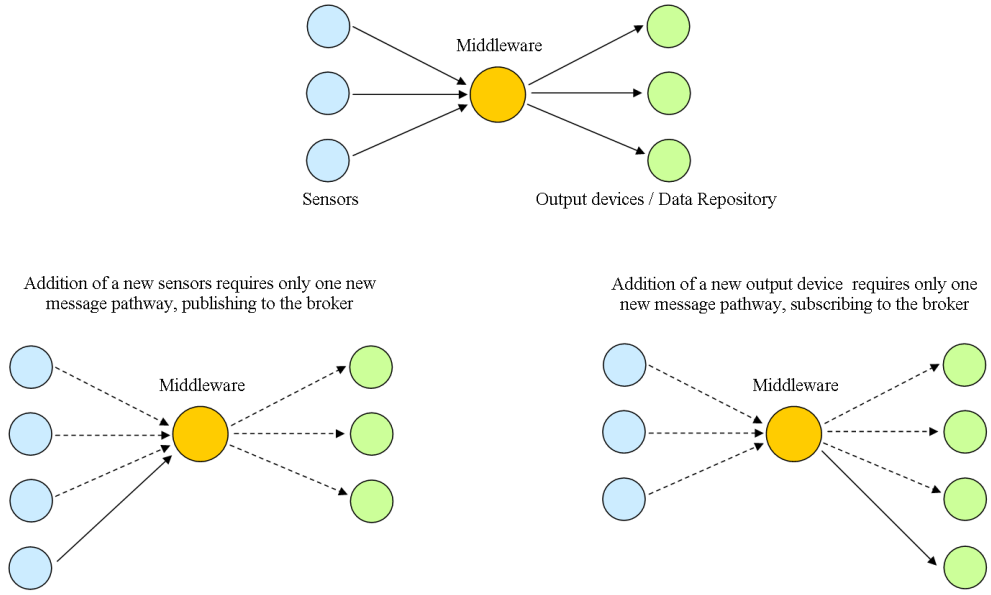


FIGURE 2.4: Addition of a new component using middleware

system much more scalable and customisable. In a traditional system, when a new component is added all the other devices would require modification. When using middleware only one new communication channel is required. This is shown in Fig. 2.3 and 2.4.

FIGURE 2.5: The Labjack UE9, a data acquisition card

2.3 Implementation

The system has been through many revisions before the current version was achieved. Originally the system was developed in Java[45], this had two major drawbacks. Firstly, this required an instance of the Java runtime environment every time a new application was run. This used up unnecessary amounts of processing power and memory load on the computer running the software, limiting the number of applications that could be launched. Secondly, issues were noticed when more than one Java application made a request to the same dynamic linked library (DLL) file, the format the Java library of the data acquisition card is available in. When two applications would query the same file the system would freeze and not respond until after a full system restart. To resolve these issues the software was developed using Perl[46].

To provide better scalability and ease of deployment, the system was built as a number of small components each subscribing and publishing data to the message broker. The system was broken down into three core components: data acquisition, calibration and storage. After initial testing of the system a fourth component was introduced, a change monitor; this will filter data that has not changed by a threshold set by the user. These stages are described below.

2.3.1 Data Acquisition Card

The data acquisition card used for a typical deployment in this system will be the Labjack UE9[47], shown in Fig. 2.5. Other options will be included to show the extensibility of the implementation.

The Labjack has 14 analogue inputs, 2 analogue outputs and 23 digital input/output (I/O) channels. The analogue inputs can be set to unipolar $0 - 5$ volts or bipolar $\pm$ 5

volts with 16-bits resolution and the analogue outputs have a range between 0 – 5 volts. The digital I/O can either be configured as input, 5 volt output, 0 volt output, timers or counters. The Labjack UE9 provides both a USB and Ethernet interface. For the environmental monitoring the Ethernet interface will be used connected directly into the local area network (LAN) of the University of Southampton. In experimental set-ups where a computer with network capabilities is already available the USB interface will be used. To make use of all the I/O channels of the Labjack UE9 an additional breakout board is required, the CB37, and is connected to the DB37 connector of the Labjack. Several software examples are provided for the Labjack UE9 covering programming languages such as Java™, LabVIEW™, MATLAB™, Visual Basic and C as well as a test/control panel for Microsoft Windows™.

2.3.2 Acquisition component

The acquisition component was developed to query the data acquisition card via a socket connection and publish the raw data to the message broker. The component is invoked with seven command line parameters, the message broker address, the topic as the location of the deployment, the address of the data acquisition card, a comma separated list of analogue sensors, a comma separated list of digital sensors and a comma separated list of acquisition channels of the sensors. When the component is launched both sensor strings and the channel string are spilt on the comma and stored in an array. The software then connects to the message broker and enters a loop where the data is published at a rate set by the user. Within the loop the software attempts to connect to the Labjack. If a connection is made the component loops through each sensor in the analogue and digital sensor arrays. For each sensor the component sends a request to the Labjack for the appropriate channel value via the socket connection and retrieves the response. This value is wrapped into an XML message in the format shown in figure 2.6 and is published on the raw data message topic. A Unix time stamp is added to the message at this point so when stored the time of acquisition is used rather than time of storage. The message pathway including acquisition is shown in Fig. 2.7.

```
<msg>
   <data> Sensor name </data>
   <value> Raw data </value>
   <unixtime> Unix time stamp </unixtime>
</msg>
```

FIGURE 2.6: XML format for raw data messages

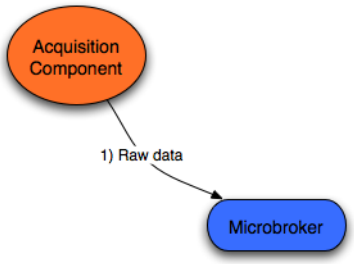


FIGURE 2.7: The messages pathway including acquisition

2.3.3 Calibration component

The calibration component was developed to retrieve raw data and produce calibrated data, acting as both a subscriber and a publisher. The component is invoked with two command line parameters, the message broker address and the topic to subscribe to. The component subscribes to a range of channels where raw data is published, using the # wild-card. When a messages is received the XML within the messages is parsed into an associative array. The sensor type is contained within the data tag; this is used to search a look-up table of calibration methods. If the sensor name is found the calibration subroutine is retrieved and used to calibrate the data. The subroutine also contains the units the data has been calibrated to. This calibrated data, along with the sensor type, is wrapped back into XML and published to a calibrated data channel, shown in Fig. 2.9. The format of the messages containing calibrated data is shown in Fig. 2.8. If a sensor name is not found in the calibration subroutine array the component publishes a message to an error log topic and the data is ignored; the system administrator can review these errors using the administrator panel.

```
<msg>
   <data> Sensor name </data>
   <value> Calibrated data </value>
   <unixtime> Unix time stamp </unixtime>
   <units> Unit of calibration </units>
</msg>
```

FIGURE 2.8: XML format for calibrated data messages

2.3.4 Change component

The change component was added after the initial testing of the system. It was decided to acquire data at a rate of once every one to three seconds. This provided sufficient resolution in the data to observe all the changes which would affect an experiment. It was noticed that typically there was very little change in data within this time-scale, leading

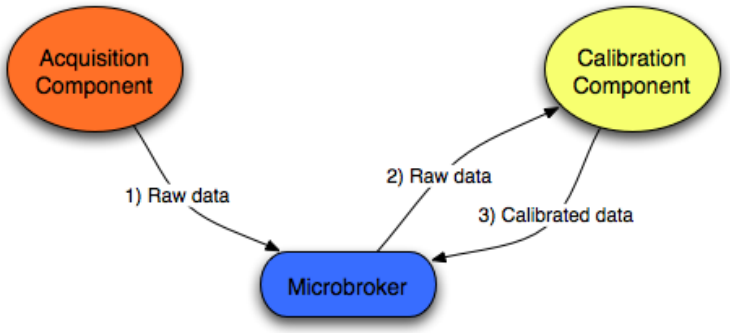


FIGURE 2.9: The messages pathway including acquisition and calibration

to a lot of repeated data in the database. To prevent the repository from becoming overloaded with data the change component was developed to filter out repeated results.

The change component is invoked with four command line parameters, the message broker address, the topic to subscribe to, a comma separated list of sensor names and a comma separated list of values used as the threshold values. The software subscribes to a range of calibrated data channels on launch. The component splits both comma separated lists from the command line and creates a associated array using the sensors names as key and the threshold values as the value, referred to as the changes array. The component also creates a second associated array using the sensor names as key and leaving the values undefined, referred to as the previous data array. When a message is received the XML is parsed into a new associated array. The message can then be handled in one of four ways.

- Sensor not recognised – The sensor name is used to search the previous data array; if the sensor is not found a message is published to the error log topic and the data is ignored.
- First piece of data from the sensor – If the sensor name is found in the previous data array but the value is undefined the calibrated data is stored in the previous data array and a new message is published on the changed data topic, the format of a 'changed data' message is shown in Fig. 2.10
- New data from the sensor within change threshold – If the sensor name is found in the previous data array and the value associated with that key is set the component determines the difference between the values by simple subtraction. The component then searches the threshold array to retrieve the threshold value for the given sensor. If the modulus of the change is less than the threshold value the new data is ignored.
- New data from the sensor outside change threshold – As with the above case a changed value is calculated between the new data and that retrieved from the

previous data array. If this change is greater than the threshold value the new data is published on the changed data topic and the previous data array value is replaced with the new data.

The way the changed value was calculated went through several phases of testing and was originally calculated as a percentage change. The testing showed that percentage changes would not show fine resolution, if required, in data at large values. To avoid this problem the software was changed to simple calibrated value changes, which allows the user to specify an exact data value for the threshold; for example a value of 0.5 referred to a temperature sensor would represent a change of 0.5 Kelvin or Celsius whereas 0.5% is ambigous. Fig. 2.11 shows in the message pathway including the change component.

```
<msg>
   <data> Sensor name </data>
   <value> Changed data </value>
   <unixtime> Unix time stamp </unixtime>
   <units> Unit of calibration </units>
</msg>
```

FIGURE 2.10: XML format for changed data messages

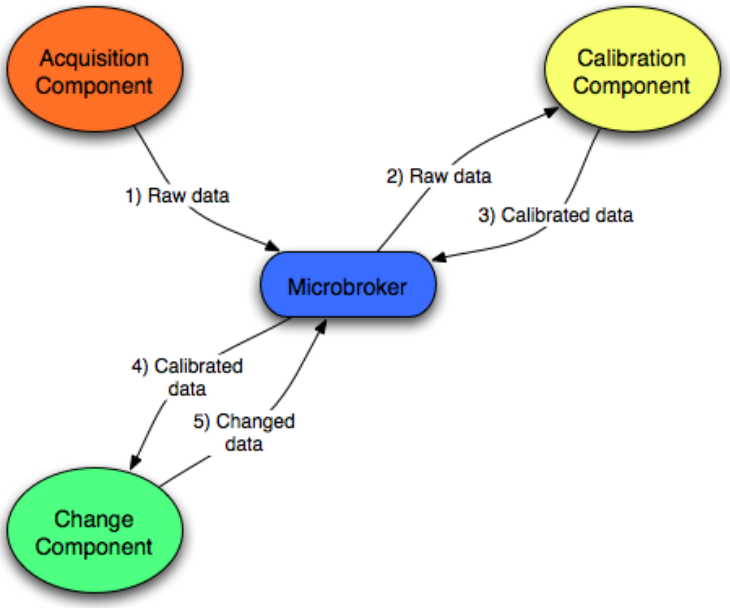


FIGURE 2.11: The messages pathway including acquisition, calibration and change detection

2.3.5 Storage component

The storage component was created to receive the data stream and store it in a logical and structured way. The component is invoked with two command line parameters, the

message broker address and the topic to subscribe to. The component can subscribe to either calibrated or change data channels, when a message is received the XML is parsed into a hash table as before. Before the data can be archived several manipulations are carried out to convert the data into a format suitable to the particular repository. As the system is designed in a modular format, multiple storage components can be run in parallel, each subscribing to the same topic of the message broker. This allows several storage components to archive the same data in different repositories, allowing comparison of these repositories. In this work a number of repository solutions have been investigated and are discussed in chapter 3. The completed message pathway is shown in Fig. 2.12.

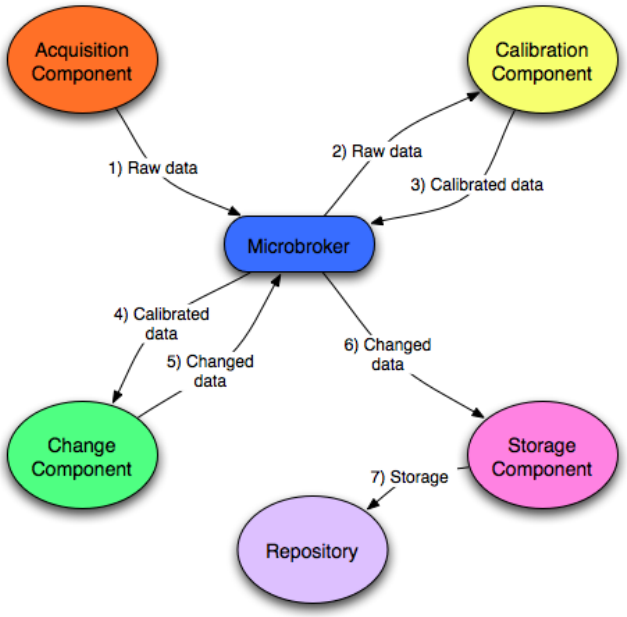


FIGURE 2.12: The messages pathway including acquisition, calibration, change detection and storage

2.3.6 Additional components

The components above describe the core data acquisition system. Additional to these, a number of secondary components were written to provide functionality requested by the users. A number of these are described below.

2.3.6.1 Alarm component

One request made by the users after the system had been running for a short while was for an alarm system. This would send an alert to the users should the values move outside of user defined limits. This was implemented as a further component in the acquisition system. The agent is launched with six command line parameters, the

message broker address, a topic to subscribe to, a comma separated list of sensors to monitor, two lists for lower and upper alarm limits and an email address for the recipient of the alarm. This component subscribes to a 'calibrated data' channel; this was chosen as the change component may filter data outside the alarm limit if the threshold had been set too high and the alarm would be missed. When an item of data is received it determines if the value is within the alarm limits. If the data is outside of the limits the components generates an email detailing the sensor, time and value detected and sends it to the recipient. The component will also send an email once the sensor values are back within the alarm limits. Once an alarm has been triggered the component will not send another alarm for 5 minutes. This prevents multiple emails being sent if the sensor reading is fluctuating at the alarm limit.

2.3.6.2 Inference component - Room occupied

By combining two or more of the primary streams of data, new 'inferred' data can be generated. This was demonstrated through the development of an 'occupied' component. This component subscribes to the passive infra-red (PIR) sensors and reed switches of a given laboratory or room. The software uses the data from these sensors to attempt to determine if the room is occupied. This is implemented by detecting the opening of the door through a trigger of the reed switch. The PIR sensors are then monitored to determine movement within the laboratory until a second trigger is detected on the reed switch. The software can also be extended to determine if it suspects someone is still in the room and no longer moving, this can then be linked to the alarm system to alert laboratory officers of potential injuries. These new 'sensors' are treated as any other, the values are published to the message broker and archived via the storage component.

2.3.6.3 Off-line data capture

As a network connection is not always available it was necessary to develop a system to allow data capture off-line. This was achieved with a modified version of the acquisition component. To allow communications with the Labjack it can either be connected through the universal serial bus (USB) port or using local network settings. When the data has been captured the MQTT message was constructed as before but instead of being published it was amended to the end of a cache file with the topic to publish it to. When communications with the Microbroker have been re-established the file is read line by line and the messages are published to the corresponding topic. As the message contains a time stamp of when it is acquired the time of collection is recorded rather than time of storage.

2.3.7 Initialiser/Re-launcher

The software was designed to run on a Linux based system running Perl. It was decided to run a unique set of the components (acquisition, calibration, change and storage) for each location monitored. This allowed for customisation of the software for particular locations, such as a lower acquisition rates, alarm systems or inferencing. This was implemented on the Linux machine as a number of GNU screen[48] sessions named as the location and the components type (E.G. PID.27:1005.acquisition). For the deployments at the University of Southampton all the components ran on a single server, although as discussed above some components could be moved to external hardware.

After initial testing it was found that network 'blips' could cause the connection to the Microbroker to be lost and the client to terminate. To overcome these uncontrollable incidents a further Perl script was developed to determine which components had terminated and relaunch them. The script queried a database containing information on the deployment, discussed further in chapter 3, to determine which locations are monitored, the sensors in that location and the configuration data related to those sensors (change thresholds, alarm limits). The software would query which screen sessions were running, searching for each components using the 'grep' command. If finding that one had terminated the start-up command was generated and the components relaunched. This script was run at start-up of the Linux machine to initially launch the acquisition system and every hour after that, using cron, as a review.

2.4 Discussion

Sensor deployments began in December 2007 in a single laboratory within the University of Southampton and several more have since been included. During this period the software has streamed and stored data with minimal down-time, capturing over 50 million data items, equating to approximately 6Gb. There were several periods of extended down-time when sensor data was not captured. After investigation it was determined that this was due to external sources, for example loss of network during scheduled upgrades or loss of power during building works. The software was designed to handle situations such as loss of power and after rebooting the data stream resumed without intervention. After the first network outage the data stream from the acquisition card did not restart automatically causing the connection to the message broker to be lost. This was due to insufficient error handling by the acquisition agent. This was overcome by modifying the components to have a shorter time-out on the data acquisition cards connection than that of the connection to the Microbroker; allowing capture and response to this error and reporting through the Microbroker. After this modification the software is robust to longer network disconnections.

The acquisition agents were run at a sample rate of once per second for digital acquisition and once every three seconds for analogue. These rates were chosen as it provided sufficient sampling to capture the slow changing environmental data of the laboratory. The sampling can run faster when required, for example capturing experimental data. The system has currently been tested up to once every millisecond with no problems.

The use of the Microbroker allowed for faster development of the basic acquisition system. Small components could be built and tested independently to the overall solution before being integrated. This is most significantly shown by the change agent. Initially the system was developed without this component but after running the system for several days it was determined that much of the data was repeated readings. The change agent was produced to avoid this redundant data capture. On integration of this component the only modification to the system was to change the topic to which the storage agent was subscribed.

As new laboratories were added to the system new instantiations of the acquisition system were launched, subscribing to the appropriate topics. This approach was used to extend the modularity of the system; each data acquisition card having an independent message pathway that can be controlled as required. The Microbroker allows this as each laboratory has its own hierarchical branch through the topics. The Microbroker can also be used to send the start and stop control messages to the appropriate components.

Although most laboratories that were monitored used the Labjack as a data acquisition card, this was not possible in the Physics femtosecond laser laboratory. This laboratory already had a Pico TC-08 serial data acquisition card with four thermocouples capturing data to the local machine. The acquisition component was developed in a format to allow simple modification to work with new data acquisition cards and therefore required only minor changes to capture this data.

The off-line data capture software was used to capture environmental sensor data while collecting naturally occurring slick samples in a RIB in the Solent. The software was run on a IBM ThinkPad with Debian distribution of Linux. On the first outing the amount of data captured was limited to the battery life, which was approximately two hours. On the second trip two batteries were taken, the first lasted for approximately two hours and the second stopped working after about five minute (it is thought the battery was not charged before leaving). A plot of the slick collection points can be seen in Fig. 2.13.

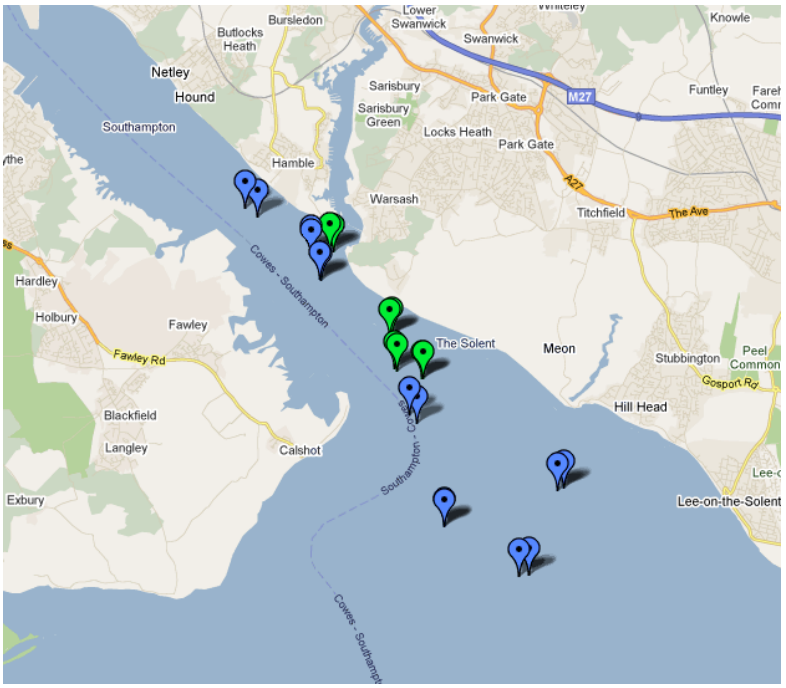


FIGURE 2.13: The slick sample locations collected while monitoring environmental conditions using the off-line data capture component on a boat in the Solent

It is planned to run the software on a more light-weight system such as the LinkSys NSLU2, shown in Fig.2.14, flashed with a modified version Debian. This could be connected directly to the 12v battery of the boat, allowing the software to run uninterrupted during the entire sampling process.

FIGURE 2.14: The LinknSys NSLU2

During the time the software has been deployed it has provided useful information on the changes of environmental conditions during experiments, helping to understand why some experiments have given unexpected results. The data was mostly used as a subsidiary check of the laboratory conditions, validating results. On a number of occasions the data has proven to be highly useful:

- Office break in – During the early stages of deployment a data acquisition card was deployed in the office, capturing temperature, humidity and light levels. One night the office was broken into, although no items were stolen. As the lighting in the office was auto-sensing the lights came on automatically. Using the data captured from the light sensor the exact time of the break in could be determined and used to search the CCTV records to identify the intruders.

- Server room air conditioning malfunction – The Frey group run a number of servers for the research projects under development. A data acquisition card was deployed in the room used to store the servers, monitoring access, movement, temperature, humidity and light levels. The alarm system was also deployed to alert if the temperature of the room increased towards a level that may damage the machines. On several occasions the air conditioning system developed a fault and the temperature increased to above the alarm levels. Once the alarm had been received a temporary solution to lower the temperature was put in place until the air conditioning was fixed.

- Femtosecond laser power drift – The power output of the femtosecond laser greatly affects the observed results. Over a number of experiments the power output was varying dramatically and the users could not determine the cause of this. Through looking at the data from the plant room and laser box it could be seen the temperature was fluctuating, resulting in the varying power output. Using this data the cause was related back to the air conditioning cooling in bursts; it is now planned to switch to a constant cooling system.

Chapter 3

Data storage

3.1 Background

One of the main aims of the data acquisition system is to be able to review historical data. This allows the environmental conditions to be considered and compared when looking back over older experimental data. As the data passing through the message broker has a short lifetime (a maximum of the last message received using the retained flag) a repository is required to archive the data. If this data is not handled correctly it will lead to a deluge of data[49], such that the tools cannot process the volume of data being generated. A solution is to include metadata, data about the data, providing provenance and context to the data. This metadata can then be used to build extended searches to extract detailed information derived from the data. Through inclusion of this metadata the integrity and trust in the data is also increased[27].

Borgman *et al.* describe a data life cycle for capturing sensor data[50]. This work identifies a number of stages a sensor deployment goes through and identifies the metadata concepts to be captured during each stage to increase the trust in the data. The key concepts relevant to this work are:

- Location - Where the sensor is deployed and how that sensors related to the overall environment
- Sensor type - The manufacturer and model number of the sensor
- Calibration - The calibration process of the given sensors and time of last calibration
- Sensitivities - The range the sensor is capable of capturing and its resolution within that range
- Measurement types - The units of the acquired data

When developing or deploying a repository solution for this work the ability to include these factors must be considered. By providing an API with access to this metadata, machine-to-machine interaction can be included. By following open standards for marking up the data the software can be integrated with external software, such as a mapping or post processing server.

As discussed in section 2.3.5, a number of data repositories were run in parallel. This allowed comparison of the benefits and disadvantages of each type of solution. A number of these repositories are discussed below.

3.2 MySQL data storage system

The initial repository solution used in this work was a MySQL database[51]. MySQL was chosen as it provides a flexible, fast and reliable system for storing the data. A custom schema was generated to hold not just the data itself but the metadata such as deployment details and calibration methods. The schema is shown in Fig. 3.1.

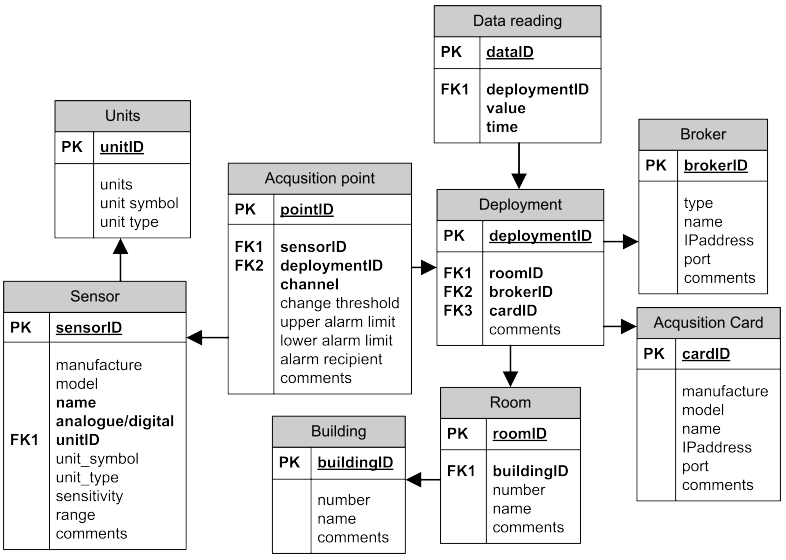


FIGURE 3.1: The MySQL database schema used for storing environmental data

This schema went through a number of revisions during the development of the system. Originally only the basic information such as the topic of message, sensor value and a time was stored. This solution stored no metadata and therefore the users had little trust in the data. This schema was expanded over a number of steps to produce the final implementation. To accompany the MySQL database and expand the available metadata, deployment schematics were generated showing sensors locations within the specific laboratories. These were made available through a web interface discussed below.

3.3 Open Geospatial Consortium Sensor Observation Service

The Open Geospatial Consortium (OGC)[52] have developed a framework of open standards for integrating sensors and sensor networks over the internet. This architecture has been named sensor web enablement (SWE)[53]. This framework describes a number of APIs for discovery, exchange and processing of the data recorded by the sensors with the network. The framework also includes a large number of specifications including observation and sensor encoding, alert systems and a sensor repository, the sensor observation service (SOS)[54]. The SOS specification describes an API for managing sensors and retrieving those sensor's data, for either dynamic or in-situ sensors. The use of the global standards described in the SWE specification will enable interaction with systems running other sections of SWE. These can be located either in the University of Southampton or any other location across the globe. The SOS specification also includes the storage of a large amount of metadata (such as sensitivities, operational ranges and global position), increasing the trust in the data.

There are three core operations the SOS must provide, getCapabilities (this retrieves all operations and allowed values the SOS offers), describeSensor (this returns the metadata of a given sensor, typically in SensorML format) and getObservation (used to recall data, typically return as SensorML) along with a number of non-mandatory functions. The SOS uses the concept that an observation is an event that produces a result, the result being an estimate of the observed phenomenon. Each event is classified by a time stamp, a feature of interest, the observed phenomenon and the procedure (sensor). Data is recalled from the SOS through observation offerings, which are non-overlapping groups of related observations. A feature of interest is an area, defined by global position, that contains one or more sensors within the sensors network[55].

In this work, an implementation of the SOS specification, developed by 52^oNorth[56], will be used. This implementation is programmed in Java and runs on a Tomcat[57] server. The solution uses PostGreSQL[58] as the back end database, although access to the data is controlled by the API. 52^oNorth have developed the API using the remote procedure call architecture[59]; all calls are made to single address and the content of the request describes which function is being called.

3.4 Other repositories

3.4.1 RRdtool

RRDtools advertises itself as an 'Open Source industry standard, high performance data logging and graphing system for time series data'[60]. This software can be used to store

a given amount of data and average it over specified time blocks (for instance a user may have three tables, one containing data for a given day and average per minute, the second containing data for a month and averaging per day and the third containing data for a year and averaging for a fortnight). The software can also be used to create plots of the data in a number of formats and output types such as PNG, JPG or directly to web pages. The raw data can also exported from the database as XML.

A limitation of this software is that it only allows data to be entered as a simple time stamp and value, therefore no metadata can be associated directly through that system. A possible solution to this is to produce a wrapper website and subsidiary database, this could be used to associate the data obtained from the RRDtool database to the metadata known about the system.

3.4.2 PaChube

PaChube[61] is an on-line sensor repository solution produced by Connected Environments. This web-base repository allows users to create feeds, such as a laboratory, which contain a number of data streams, such as sensors. The feed is updated using a representational state transfer (REST) interface where data can be entered through HTTP PUT requests as either comma separated variables (CSV) or extended environments markup language (EEML)[62]. Feeds are displayed to the user on a Google Maps interface where markers represent a feed, and therefore each feed has a GPS position associated with it. When a marker is selected the user can view a pop-up containing feed information and a 24 hour plot for each data stream. Feeds contain the previous 24 hours data in 15 minute time segments and can be downloaded in a number of format (CSV, XML or json). The PaChube website also provides several user generated add-ons which present the data in a number of different ways. As the data is stored remotely the users have no control over the data policy. This means data may be lost through changes in the storage policy, implementation of the system or discontinuation of support.

3.5 Web interface

A web interface was developed to request data from the MySQL database and present it in a number of formats. The web pages were developed in PHP: hypertext preprocessor (PHP) running on an Apache2[63] web server. As well as the data monitoring pages a number of additional features were added. A user could review the details for the deployments including the change thresholds, upper and lower alarm limits and additional comments relating the individual sensors. Deployment schematics for each laboratory

were also generated and available through the website. These diagrams show the extract location of the sensors within the laboratory, providing some of the recommended metadata described above.

3.5.1 Data monitoring

To allow real-time monitoring an additional acquisition component was developed to store the most recent values from each sensor into a database table. A database was created with a single table containing three fields, a location, the sensor information and the most recent value. The software subscribes to the calibrated data from a single location and when a message is received it is parsed as before. The database table is searched for a record containing the same location and sensor information and if not found the data is inserted. If a record is found the current record is updated with the new sensor value.

A web page was developed to display this information and update regularly so a user may track changes to the environmental conditions. The page first displays an entry box where the user selects the sensor or location they wish to view; % is a wild-card used to search for all entries containing that value. On submission the page will display all the most recent values found for that location and update every 5 seconds to show any changes that have occured; an exmaple of this is shown in Fig. 3.2.

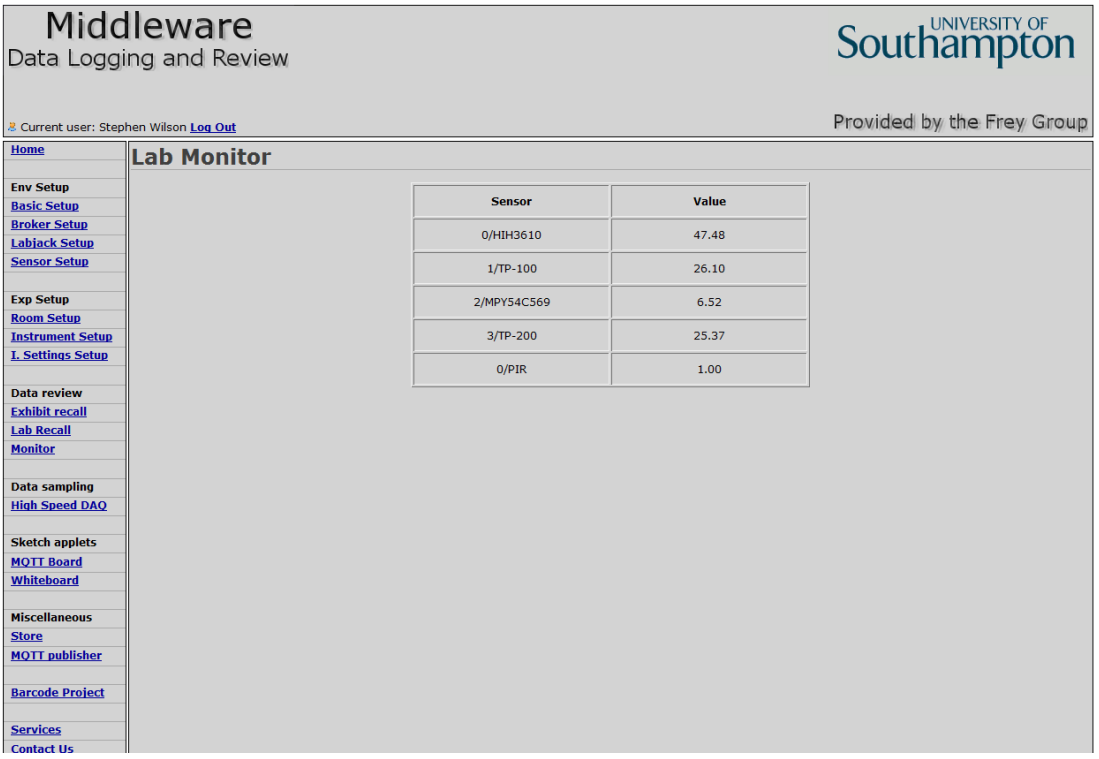


FIGURE 3.2: The real-time data monitoring web page

3.5.2 Data recall with Simile TimePlot

A system to allow graphical representations of the data was also developed within the website framework. After investigation of the available plotting software it was decided to use the Simile TimePlot software[64]. This produces a graphical representation of the data and provides details of data points on mouse over events. The Simile TimePlot is generated from a number of comma separated text files and XML documents containing the sensor data.

When wishing to recall data in a graphical format the user must first enter the location as before as well as a start and end time in the format, YYYY-MM-DD HH:MM:SS. The user input will be entered as local time and therefore must be converted into Greenwich mean time (GMT) before used to query the database for sensor data within this criteria. A new text file is generated for each analogue sensor returned by the database query and the data related to the sensor written to that file; a new line is written for each piece of sensor data in the format *time,value*. The returned results are in GMT and therefore must be converted back to the local time zone before being written to the file. The digital sensor data is stored as an XML file (one for each sensor). The format for these files is shown in figure 3.3. A system to allow a data series to be hidden and re-shown was integrated; when the TimePlot is generated a check box referenced by the sensor information is also created. When a check box is selected a JavaScript function reloads the TimePlot removing or adding the selected data. An example of this is shown in Fig. 3.4.

```
<data>
   <event start="MM DD YYYY HH:MM:SS GMT" title="Topic of sensor data">
   </event>
   <event start="MM DD YYYY HH:MM:SS GMT" title="Topic of sensor data">
   </event>
</data>
```

FIGURE 3.3: XML format for the Simile TimePlot digital sensor data

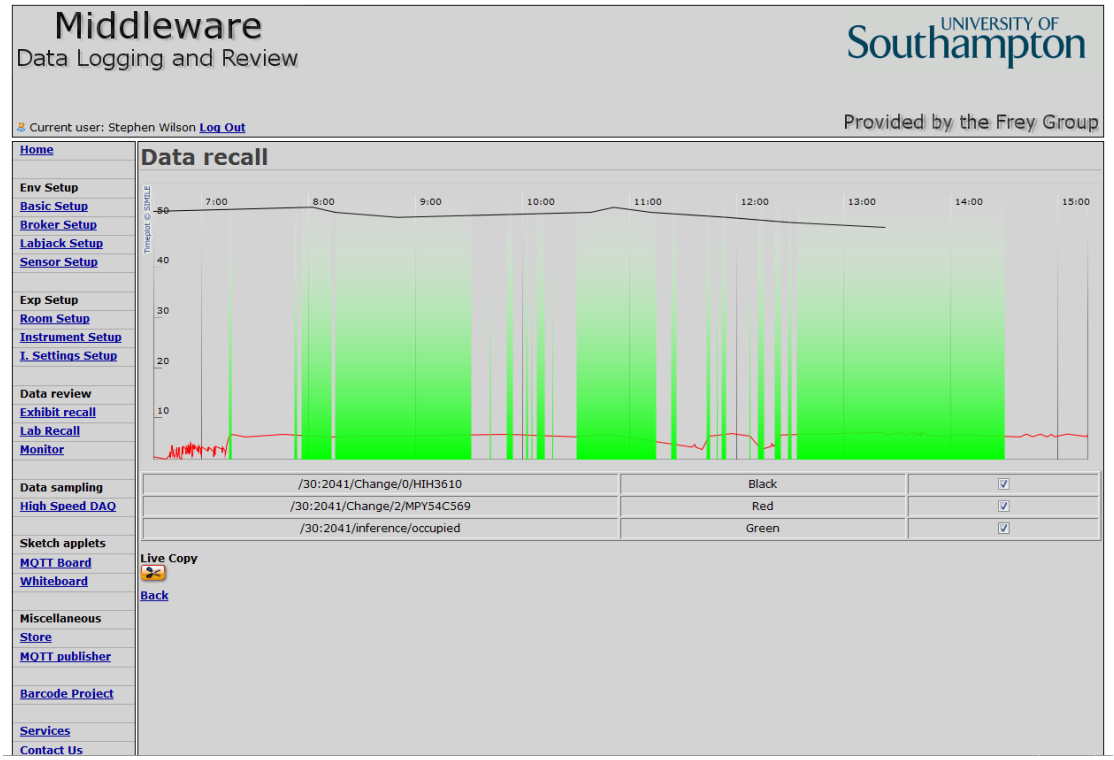


FIGURE 3.4: The Simile TimePlot data recall web page

3.5.3 End of day blog entries

As part of another project, a blog based electronic notebook is under development[19]. The data acquisition system was modified to integrate with this system. A script was developed to create an 'end of day' blog post for each laboratory within the deployment database. These blog posts contain a Simile TimePlot graph showing all sensor data captured within a 24 hour period from the given laboratory. As well as creating an individual blog post for each laboratory, a summary post is also created containing details on number of sensor readings stored per laboratory on the given day. This can be used as both a maintenance tool to check data is still being collected from the required laboratories and a validation tool to check the number of points is similar to the average collected (and as such no significant changes in the environmental conditions have occurred). An example is shown in Fig. 3.5.

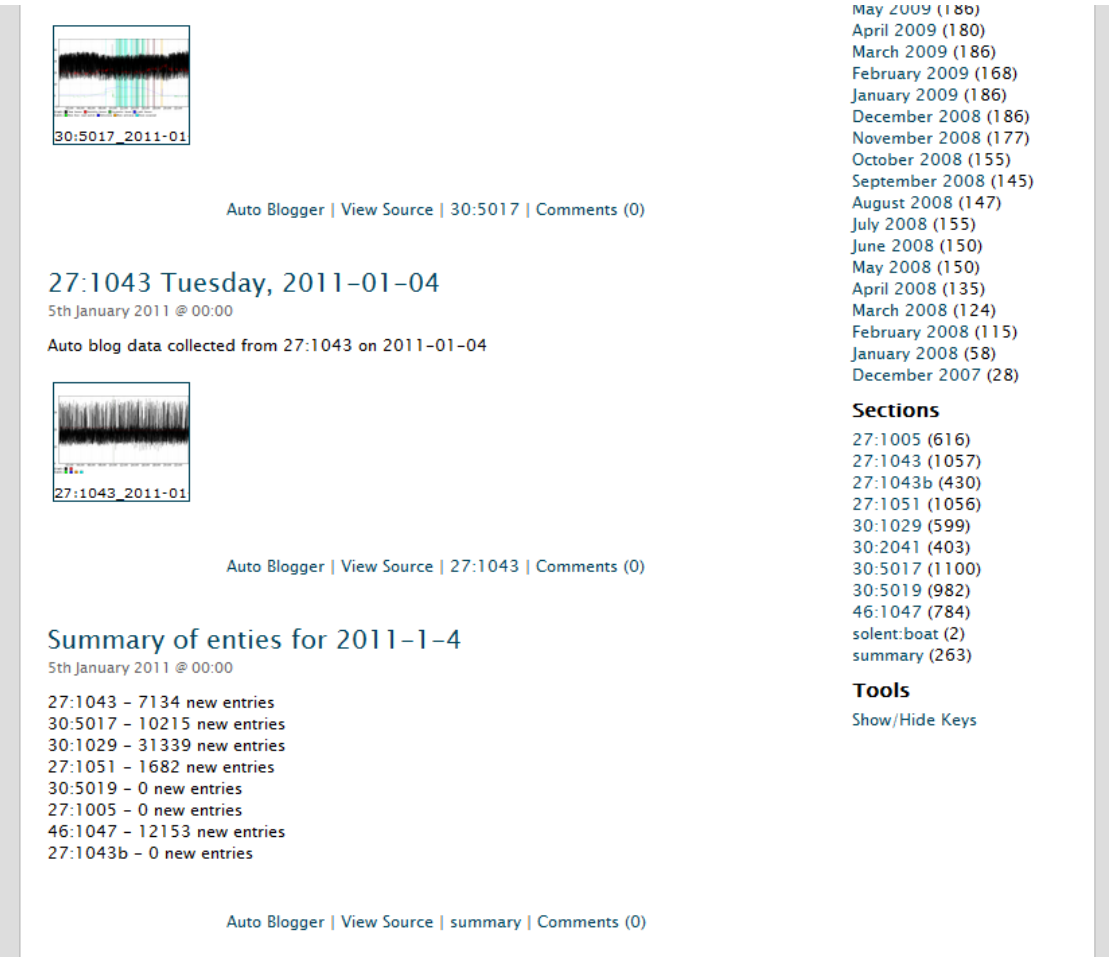


FIGURE 3.5: The end of day and summary blog posts automatically generated as part of the acquisition system

3.6 Discussion

A number of potential repository solutions have been discussed to capture data generated by the data acquisition system. The use of metadata has been discussed as a way to improve the trust and quality of the data as well as allowing more meaningful searches.

The MySQL database allows for complete customisation in the schema and can incorporate as much or little of the metadata as required. Once the schema has been implemented it is difficult to modified or extend. The MySQL solution does not provide an API to access the data as standard although one could be developed in a language such as PHP. The use of a custom schema can make it difficult to comply with the global standards such as SensorML and therefore could make it difficult to share the data with external software. The use of open source tools such as Simile TimePlot allows for rapid development of a web site for displaying the data.

The OGC SOS provides a extensive software package for capturing sensor readings and the associated metadata. The service is part of a larger specification describing a number of interfaces for capturing, displaying and manipulating this data. The specification stipulates the use of SensorML and Observation & Measurements standards to describe the data, enabling operability between each component. The 52^o North implementation of the standard uses a remote procedure call API to deposit and recall data from the archive using these standards. With this being a full developed system, following the OGC specification, for a wide range of sensors and deployment types the metadata required is beyond that needed for a laboratory monitoring system. Although this may be useful in the long term, as more metadata will be available, it make make the addition of new features and sensors a laborious process.

On-line solutions such as PaChube, discussed above, SensorBase[65] and SensorPedia[66] allow the storage of sensor data on a remote server through a web interface. These systems remove the need to administer and maintain the system as it is handled by the software hosting group. Each of these sites promote the sharing and reuse of data through a social community system developed as part of the web interface. The data is uploaded and retrieved via a standardised format developed as part of the individual project. As each solution uses its own formatting, it can make it difficult to share the data between external software without an additional step to convert the data into the required format. Each site offers the ability to capture a level of metadata, although typically not all that has been suggested above as important for this data acquisition system. There are also inherent risks in passing the storage to an external source. The lifetime of the data is undefined as the administrators may remove or modify the system at any time, losing the historic data. This is a major factor in PaChube where data is only stored for a maximum of 24 hours and with a resolution of 15 minutes. Issues may also arise in ownership of the data should it be of value, although this should be addressed in the terms and conditions of the web site. There is also the possibility that support for the site will be dropped.

The MySQL database has been running as the primary repository solution for the majority of this project. During this time it has collected over 50 million sensors readings from around 30 sensors, total approximately 4.5 GB. The schema described above has allowed the capture of metadata to be able to describe the sensor data to a level required in this data acquisition system.

Chapter 4

Experimental control and automation

4.1 Background

Since the first development of actuators and electronic motors, the way research is carried out has dramatically changed. Devices such as these have been integrated into almost every piece of the hardware used within the laboratory. The automation of this apparatus has allowed more complex experiments to be completed and in a much shorter time scale. These developments have also allowed experiments to be completed that previously were thought impossible. Implementation of tools to remotely control this apparatus has allowed researchers to carry out experiments while away from the laboratory. This was first shown in 1994 when a web interface was developed for a telescope[67]. These techniques have now been used to control a wide range of experiments ranging from a simple PID servo[68] to a Tokamak[69] and an x-ray diffractometer[70].

Through automation of the apparatus, the reliability and repeatability of the experiment can be greatly improved. Laborious tasks such as repeated positioning of components within the experiment can be hidden from the researcher. This removes the random human error from the results and also provides, assuming the correct equipment is chosen, a higher level of accuracy. The use of remote control allows the researcher to spend less time in the laboratory. This can be beneficial for a number of reason including safety issues arising from risks within the experimental environment and accuracy issues caused by the researcher being present. This also allows more time to be spent on analysis and processing of the data, rather than running the experiment. The drawback to this abstraction from the experiment is that the researcher is unaware of what is occurring during the experimental process, such as parameter values, experimental state or current results. Therefore, additional tools must be put in place to allow real-time monitoring of the experiment.

There are also inherent risks with allowing remote control over the equipment. When dangerous equipment is being controlled security implications must be considered. If an unauthorised user gains control over the device they can cause harm to both the equipment and any individuals in the vicinity of the apparatus. The security level can be applied at either the control interface level using a username/password system or directly on the message broker using encryption to encode the messages[71]. Even an authorised user may damage the equipment if correct operational limits are not enforced at the software level, such as a motor moving past a limit and colliding with another section of the apparatus. Another consideration is who is responsible for administering the system, if the apparatus is not used by the hosting company the system may no longer be maintained.

4.1.1 Related work

The most commonly used product in equipment automation at the time of writing is LabVIEW[29] by National Instruments. This software allows the creation of virtual instruments (VI) used to control the equipment and handle data acquisition via a graphical programming language, G[72]. A user creates a virtual instrument by dragging graphical representations of the instruments into the programming area. A GUI is generated in parallel with the virtual instrument; tools to include buttons, dials and plots are available through the software. The input and outputs of the individual components are connected together using 'wire', which handles the flow of parameters and variables. As this is the most commonly used software for automation there is a wide range of support for interfacing with experimental hardware. The drawback to this software is that it does not provide good support for remote operation as typically the VI will run on a computer connected directly to the instrument via a bespoke controller card or the computers serial port. The operator then controls the equipment either directly from that computer or through a shared desktop using VNC[30] software.

A number of projects are under development to provide access to instrumentation, both real and virtual, over the internet. These include CIMA[73], GridCC[74], LabNet[75] and PEARL[76]. Common Instrument Middleware Architecture (CIMA) and GridCC use an extension of the Open Grid Services Architecture (OGSA)[77] and Common Component Architecture (CCA)[78] standards to describe an instrument as a Grid service[73]. These standards describe the functions of instrument similarly to that of the Web Services Description Language (WSDL) description of web services. This service based method allows the system to be used remotely and independent of the platform and programming language making a request. The use of the Grid can also provide additional benefits such as handling the security and queuing aspects of the system. A number of other projects have looked at making experiments directly available through a web interface. HTML

forms and buttons are used to submit parameters to the apparatus and is executed by the equipment. The results are returned via the HTTP response.

4.2 Implementation

The experiment employed in this work is a SHG experiment studying the liquid/air interface. The experiment is carried out by directing a high intensity laser source onto the surface of a solution at a controlled input polarisation and the second harmonic produced at the sample is detected by a photon multiplier tube (PMT) at a controlled output polarisation. The signal from the PMT passes through a boxcar integrator and is captured by a USB enabled data acquisition card. The experiment process is discussed in detail in chapter 6.

The automation system was developed to control a number of apparatus within the SHG laboratory, including the laser source, shutter, half waveplate rotators and the experimental data acquisition card. Each piece of apparatus has a communication port, such as RS232[79], which was connected to a controlling computer. The automation software uses a centralised message broker to mediate the flow of message between control interface and equipment, as with the environmental data acquisition system.

A generic Perl package was written to handle the message broker functions, such as connecting, subscribing and receiving messages. A device specific script was written for each item of equipment in the laboratory, implementing the broker package. These scripts are available in appendix A. Each controller script contains the connection information required for the given device, such as baudrate, parity, stopbits, etc.. for RS232. Once loaded the script uses this information to create a socket connection to the device. Once connected to the device the script waits for a control message; the format for these messages is shown in Fig. 4.1.

```
<msg>
   <device> System device name </device>
   <control> control command </control>
   <value> command parameter value </value>
</msg>
```

FIGURE 4.1: The XML format for control command messages

The system was developed to take control commands in human readable format, such as 'laser on' and 'close shutter'. This design was used to aid users as well as debugging and maintenance of the system. When an recognised device command is received, the script uses a look up table to convert the human readable command into the device specific syntax; any parameters are encoded at this point. This command is then executed on the apparatus. Where possible the script would monitor the execution of the command,

such as requesting the current position of an actuator as it moves, the information is published to the users via the message broker. Once the execution has completed, a message to sent back to the broker on a response topic, again using a human readable response. Should the execution have failed, this is also published back to the message broker. The format of the response message is shown in Fig. 4.2.

```
<msg>
   <device> System device name </device>
   <status> active / complete / failed </status>
   <response> response command </response>
</msg>
```

FIGURE 4.2: The XML format for control response messages

The software controlling the data acquisition card was an extension of the standard script, which included the feature to upload the captured data directly to a database. During the development stages the options to either use the message broker or direct storage for data capture were considered. Initial testing suggested that if all laser shots from a given run (up to 4000 data points) where published, at once, to the message broker, the system may slow down and affect the entire system (including the environmental monitoring system). Therefore the Harvard architecture of direct storage was applied.

An overview of the control system is shown in Fig. 4.3.

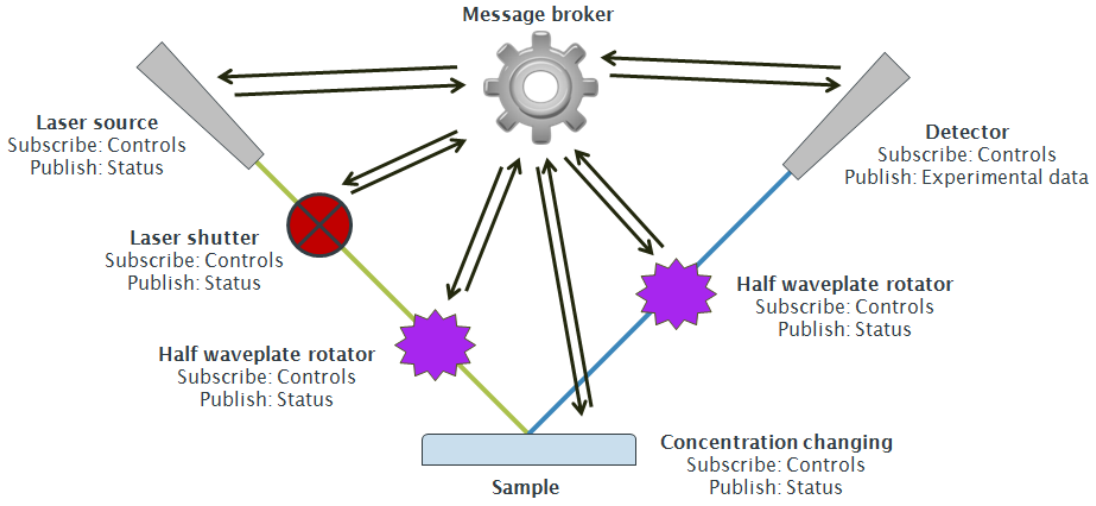


FIGURE 4.3: The message pathway for control components in the SHG experimental set up

An additional 'experimental overview' script was written to control the order of enactment during an experiment. This script used a stateful approach to order the execution of commands of each piece of equipment; awaiting a 'completed' message from the current apparatus before entering the next state. The script followed the workflows described in sections 6.5.2.1 and 6.5.2.2.

An experiment is initiated by publishing a message to the topic the 'experimental overview' software is subscribed to. This XML message contained a number of run elements, detailing the parameters for the given run (number of laser shots to capture, input half waveplate position, output half waveplate position, angle of incidence and the monochromator wavelength) as well as additional metadata describing the experiment including operator, laser source, electrical configuration, compound under investigation and experiment type. An example message is shown in Fig. 4.4.

```
<msg>
<title>SHG of COMPOUND</title>
<operatorID>1</operatorID>
<laserID>1</laserID>
<boxcarID>1</boxcarID>
<experimentType>1</experimentType>
<compound>COMPOUND</compound>
<runCount>-1</runCount>
<shutdownLaser>0</shutdownLaser>
<experimentData>
        <run>
                <runNumber>1</runNumber>
                <sampleID>1</sampleID>
                <inputPol>45</inputPol>
                <outputPol>0</outputPol>
                <laserShots>667</laserShots>
                <incidentAngle>60</incidentAngle>
                <monochromatorWavelength>266</monochromatorWavelength>
        </run>
        <run>
                <runNumber>2</runNumber>
                <sampleID>1</sampleID>
                <inputPol>12.61</inputPol>
                <outputPol>16.33</outputPol>
                <laserShots>250</laserShots>
                <incidentAngle>60</incidentAngle>
                <monochromatorWavelength>266</monochromatorWavelength>
        </run>
        <run>
                <runNumber>3</runNumber>
                <sampleID>1</sampleID>
                <inputPol>0</inputPol>
                <outputPol>0</outputPol>
                <laserShots>667</laserShots>
                <incidentAngle>60</incidentAngle>
                <monochromatorWavelength>266</monochromatorWavelength>
        </run>
        <run>
                <runNumber>4</runNumber>
                <sampleID>1</sampleID>
                <inputPol>79</inputPol>
                <outputPol>22.63</outputPol>
                <laserShots>833</laserShots>
                <incidentAngle>60</incidentAngle>
                <monochromatorWavelength>266</monochromatorWavelength>
        </run>
</experimentData>
</msg>
```

FIGURE 4.4: An example XML message used to run a SHG experiment via the overview script

As discussed above, the experimental data is stored directly to a database. The database chosen was a MySQL relational database running on a Linux server. The schema was designed to contain both the experimental results and the associated metadata. The schema used for this database is shown in Fig. 4.5. This data is made available via a web interface, described in chapter 5.

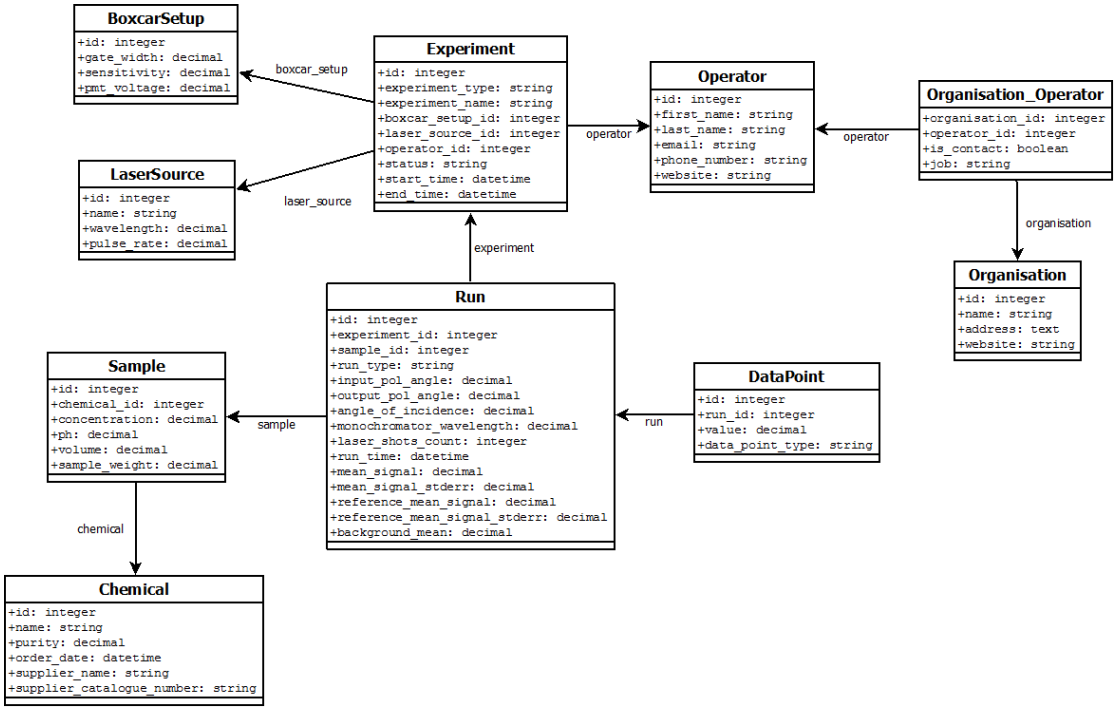


FIGURE 4.5: The database schema for SHG experimental data

4.3 Control interfaces

As each component within the control system is accessed through the message broker, the equipment can be controlled by anything that can create and publishing an MQTT message. This was expanded further through the development of a web server which publishes messages generated from a HTTP request. This allows anything that can make a HTTP request to control the equipment. A number of control interfaces have been developed and discussed below.

4.3.1 Web interface

Discussed in chapter 5 is a web interface for reviewing real-time and historical SHG experimental data and its associated metadata. An additional web page was added to the site to allow a user to control the experiment. A web form is generated containing drop down boxes to select the metadata fields, auto filled from the database. The user then enters the run parameters required for the given experiment type and submits the

form. This data is converted to the message format required by the 'overview' component and the message published. The browser is redirected to show the real-time values from the active experiment. Fig. 4.6 shows an example of the web form.

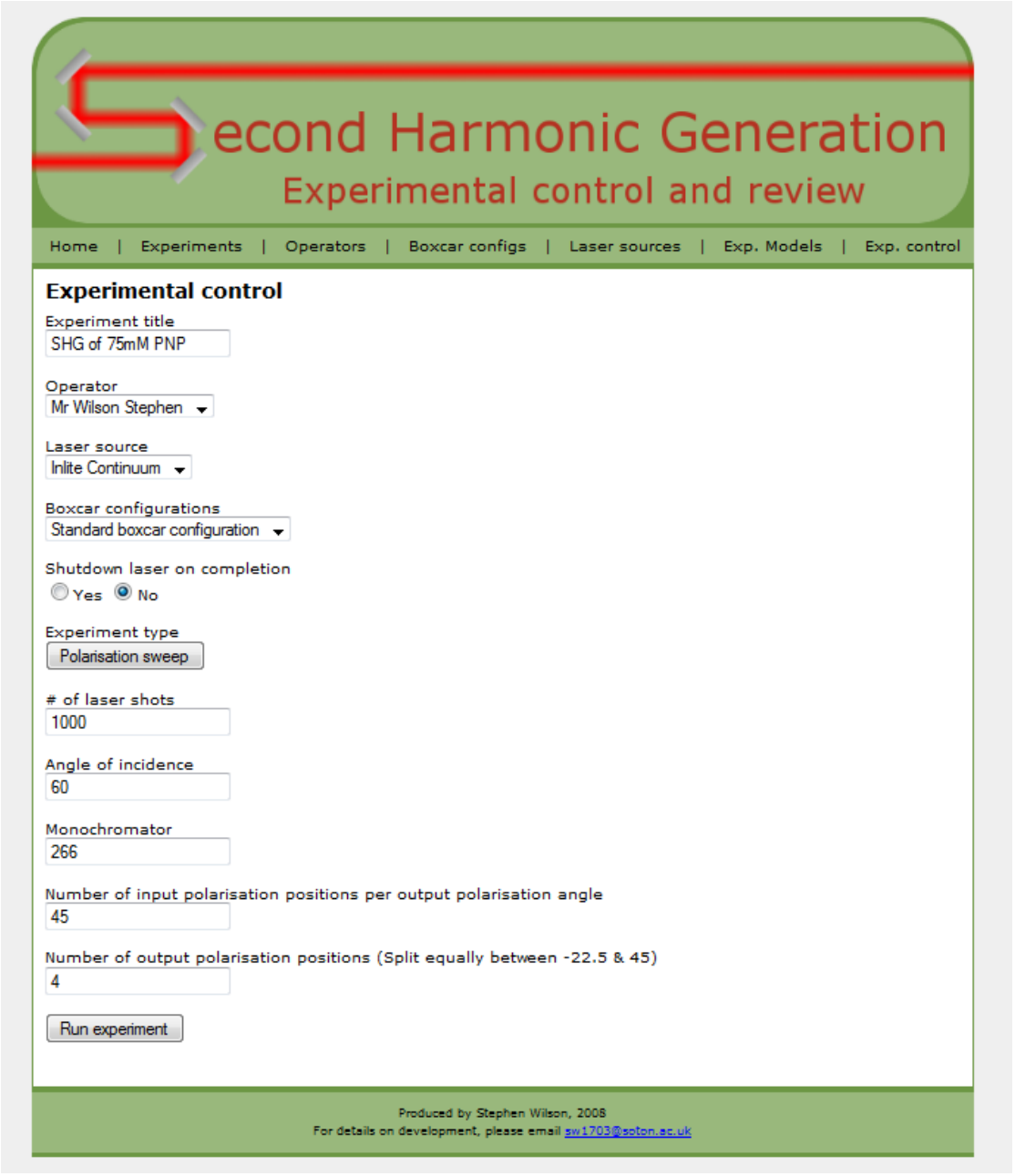


FIGURE 4.6: The web page control interface for the SHG experiment

4.3.2 iPhone application

An iPhone interface to the apparatus was developed using the iUI[80] framework. This package provides a framework for developing web pages using an iPhone like 'look and feel' while providing JavaScript functions to handle the changes in screen orientation. The application provides a menu of available devices within the laboratory, when selected

the commands available to that device are shown as buttons, sliders and text input fields. When a command is executed an MQTT message is published to the specified apparatus. Fig. 4.7 shows the home and laser source menus.

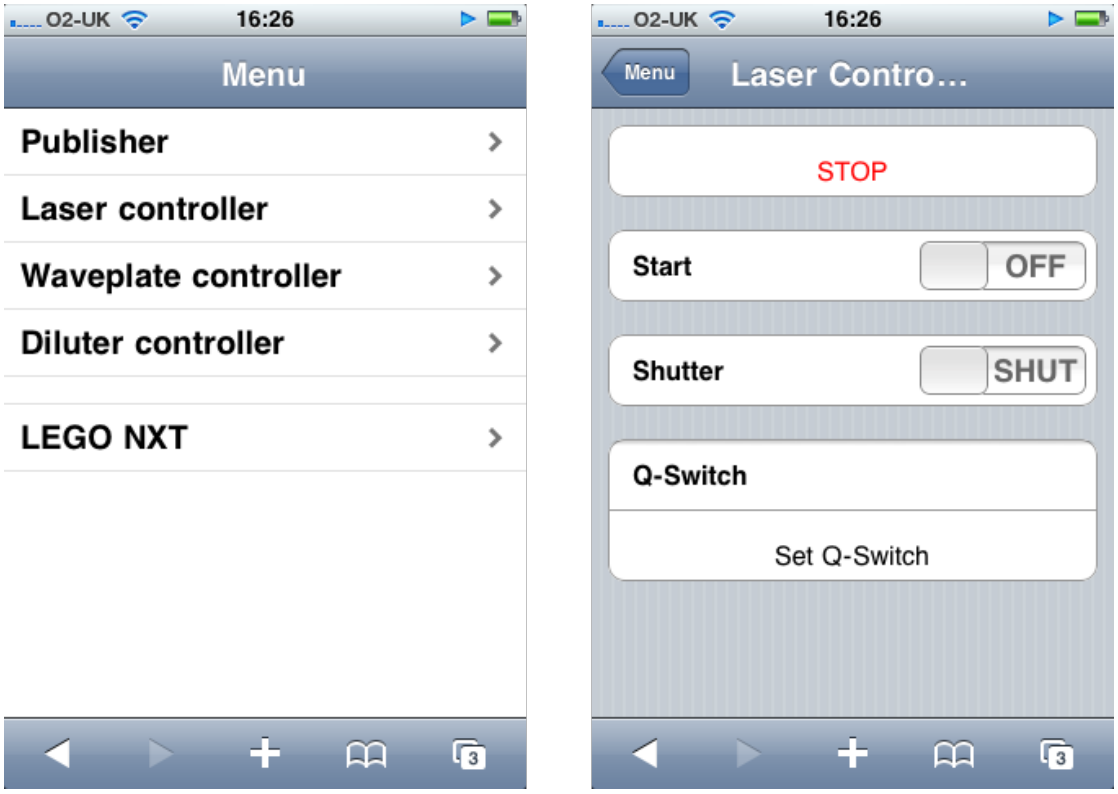


FIGURE 4.7: The iPhone control interface for the SHG experimental apparatus. Left: The list of available equipment to be controlled. Right: The laser source specific interface

4.3.3 Virtual worlds - Second Life

A virtual world is a computer based environment, typically in three dimensions, where a person can interact and manipulate objects and communicate with others. Users are represented in the virtual world as avatars, although these are typically 'humanoid', they can have any shape and size. Virtual worlds can benefit the research sector as they offer visualisation not available through traditional simulation techniques and can promote discussion among users who are located across the globe. Second Life[81] has become the most popular of these virtual worlds with over 1.3 million users. Its success has come from its easy to use interface, global media coverage and its free-to-use policy. Second Life also allows users, assuming you have the correct permissions, to build objects within its environment and develop scripts to run within them. It is this functionality that is used to develop the SHG control interface.

A representation of the SHG experimental set up was built within Second Life. Accompanying this was a control panel with a number of buttons representing the different

controls available for the equipment. This is shown in Fig. 4.8. When an avatar 'touches', through a mouse click, a button, a HTTP request is made to the MQTT web service via a script running on the button. The web service extracts the message and topic from the URL and publishes this to the message broker. Where a parameter is required the avatar 'talks', through keyboard entry, to the button telling it the required value. Second Life also allows display of web sites and streamed media in-world. This was used to show a web cam of the SHG laboratory, providing users real world feedback on the actions undertaken in the virtual world.

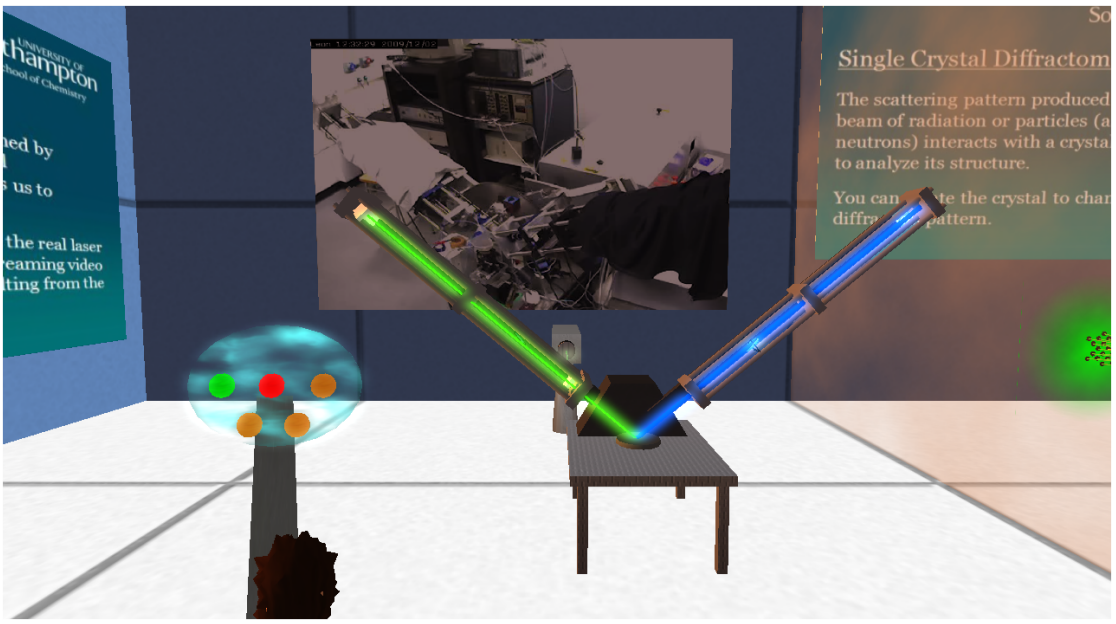


FIGURE 4.8: The Second Life representation of the SHG experimental set up with control panel

4.4 Discussion

An automated remote control system has been developed and applied to the SHG experiment. The software was developed to allow quick implementation of additional component through use of the message broker. To introduce a new piece of equipment a short device specific script is required containing the expected commands for the device in a human readable format and the mapping from those commands to the device specific syntax along with the communications protocol. An overview script has been developed to handle the completion of an entire experiment through a single MQTT message, which controls the order of execution of each component in the system and handles any errors which may occur. Experimental results are uploaded directly to a database rather than published to the message broker, reducing the load and preventing this from becoming a bottleneck in the system.

As each piece of equipment, as well as the entire experiment, is addressable via a unique topic on the message broker, a number of control interfaces can and have been developed. The integration with browsers and virtual worlds, such as Second Life, has allowed for remote access from any location and a unique visualisation of the equipment. The use of a web camera provides users with direct feedback on the actions they are performing. Through linking the remote control and real-time monitoring systems, the software allows collaboration between researchers in different institutes or locations. This can be enhanced through use of LIMS and ELN systems, providing each researcher access to previous experimental results and work processes.

To integrate a device into the system it must have a communications port, such as RS-232[79] or USB. The protocol for this port and syntax of each command must be well documented and available; this is not common practise. Typically the device is controlled by custom software, which must be purchased with the equipment. Therefore, the manufactures do not publish the required documentation to generate custom software. This documentation may also no longer be available for older equipment.

Through automation, the quality and trust in the experimental results has been greatly improved. The use of high precision automated components to control aspects of the experiment has not just removed any human error from the system but also allowed the parameters to be more accurately determined. Providing the environmental conditions along with the metadata of the equipment, such as tolerances and calibrations, the research has a much greater volume of data available to support the results. With all this extra data available, the researcher may become overloaded, therefore this must be stored and presented in a structured manner. The remote control software has also increased the safety of the experiment, the researcher can run the experiment without being exposed to the risks of the experimental environment.

In the SHG experiment security is handled through a username and password log in to the web interface, either browser or iPhone. This restricts access to the interfaces that publish the messages to the broker. This does not prevent a malicious user from snooping message published to the broker. This can be used to determine both the topics the control software is subscribed to and the format of the messages. This could be overcome by providing an encryption on the messages, preventing the user from determining the format of the messages. The Second Life control interface was only active for a short time as part of a demonstration and as such no security measures were implemented. Should the deployment have remained active, security would have been handled via an allowed avatar list. This list can either stored within a script running on the object within Second Life or in a database accessed through a web request.

4.4.1 Further work

At present, the experiment is only controlled via hard-coded interfaces. This can be improved through changing the control scripts to describe the commands they understand. This can be used to develop self-describing interfaces to extend the functionality of the middleware approach. This would be handled through a 'getCapabilities' like request to all devices of a given hierarchical level of the message broker (corresponding to a given laboratory). Each device subscribed to that topic level would send a response describing the functions it can execute, the required parameters and the limits of those parameters. The interface software can then parse the responses to generate the control interface. Through this method the interface is disassociated from the devices, allowing additional devices to be integrated without having to re-write the control interface.

The experiment controls can also be described using web or grid service standards such as WSDL[82] or OGSA[77], therefore creating an API for the experiment. The experiment then becomes a first-class object, allowing integration with external software. By integrating with the Grid, the authorisation and access control, including queuing, is handled by external software.

Chapter 5

Experimental website - Linking it all together

5.1 Introduction

As discussed in the previous chapters, the data acquisition system and remote automation software will improve the quality and trust in the data through improved experimental techniques and additional metadata capture. To maximise this potential a central point to collate all this information is required, this will be developed as a website. This will make use of existing open source tools, such as the Simile TimePlot[83] software discussed in chapter 3, to maximise its impact at minimal development costs. This will be developed in such a way to allow users to either recall completed experiments or monitor in real-time active experiments and the associated metadata.

5.2 Implementation

The website was developed in PHP[84] running on an Apache2 web server. A REST like approach was taken when developing the site, such that each resource has a unique identity (represented as a URL). The website was split into five sections: a home page, the experimental time line and results, the experimental metadata, experimental fitting results. Several aspects of the site will be discussed in detail below. Each experiment has a unique URL and these pages are split into four sections: the metadata links, a plot of the results, a downloadable table of the results and the environmental conditions. This is shown in Fig. 5.1.

Each experiment is a collection of runs. Each run is also addressable and contains three sections: a summary of run parameters and results, a plot of the run data and a downloadable table of results and background results. This is shown in Fig. 5.2.

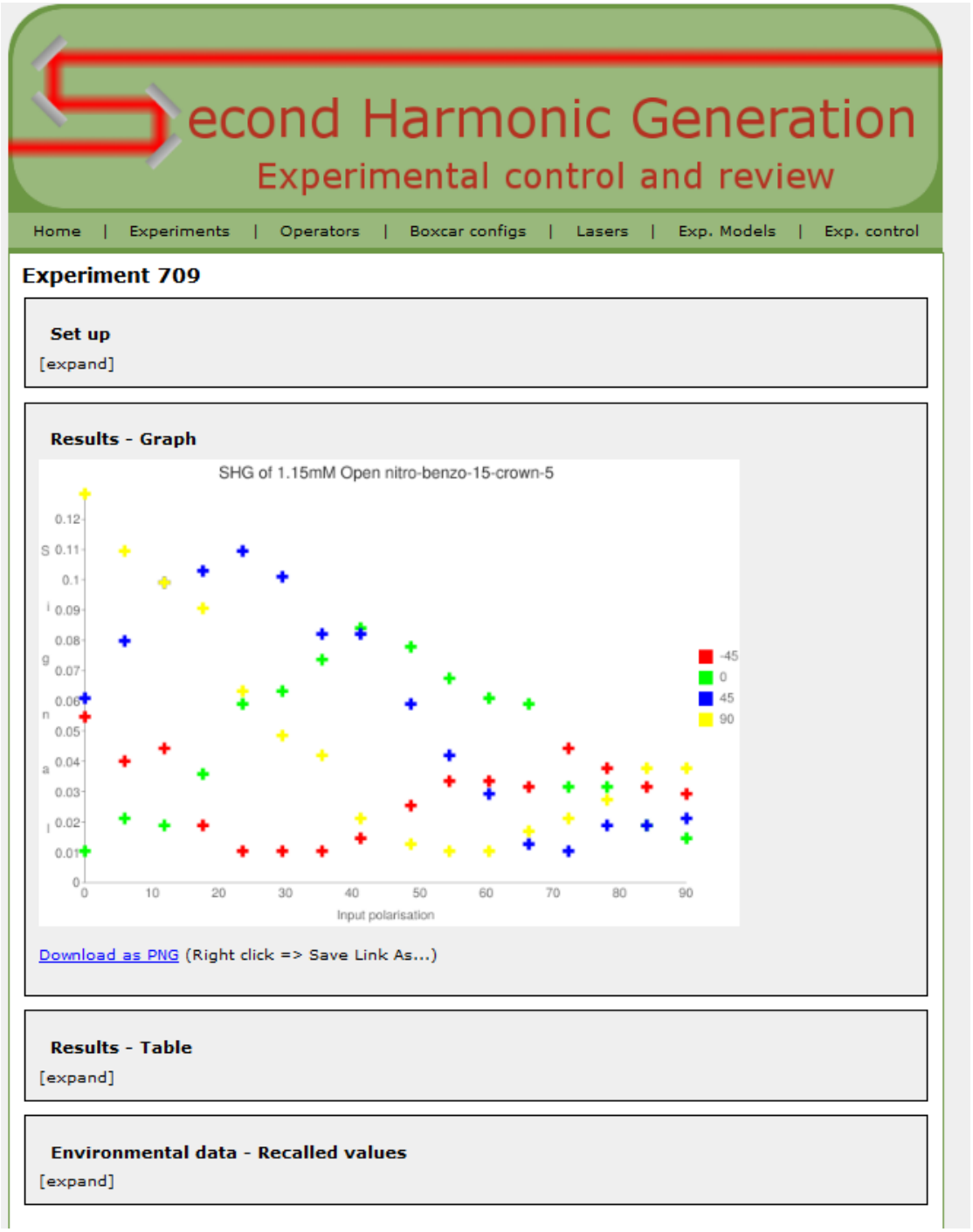


FIGURE 5.1: The SHG experimental results web page

Finally, each run is a collection of individual laser points. Each captured laser shot also has a unique URL, displaying its intensity.

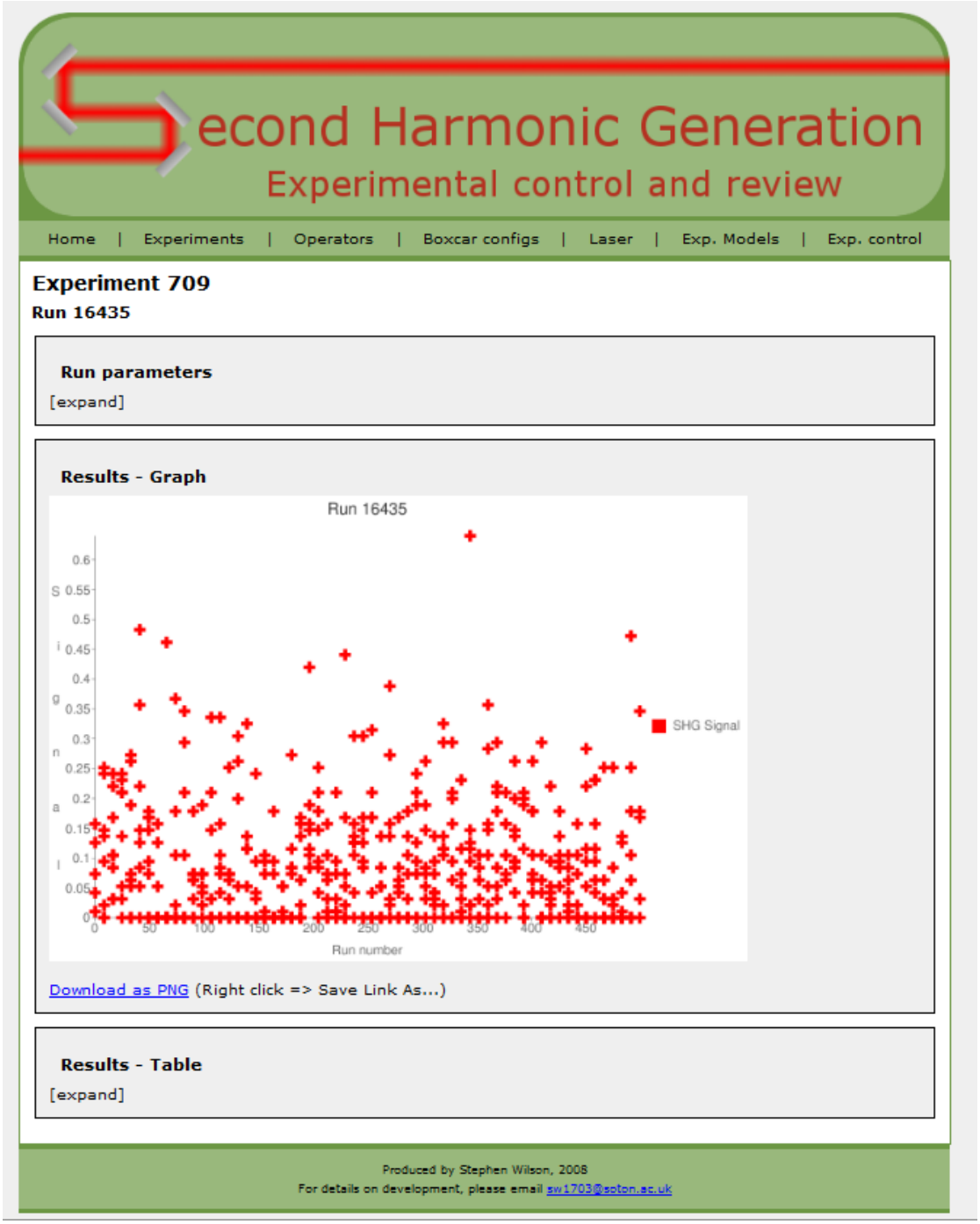


FIGURE 5.2: The SHG experimental run web page

5.2.1 Experimental Timeline

On navigating to the experimental results section a time line of results is generated using the Simile Timeline software[85], shown in Fig. 5.3. This software takes a JSON[86] representation of the experimental data and displays it as a scrollable time series. This is split into 3 sections, each at a different time resolution (hours, days and months). The user can scroll along the time series to locate the experiment they require. On clicking

on an experiment a pop-up containing a link to the experiment and a mini plot of the results is shown.

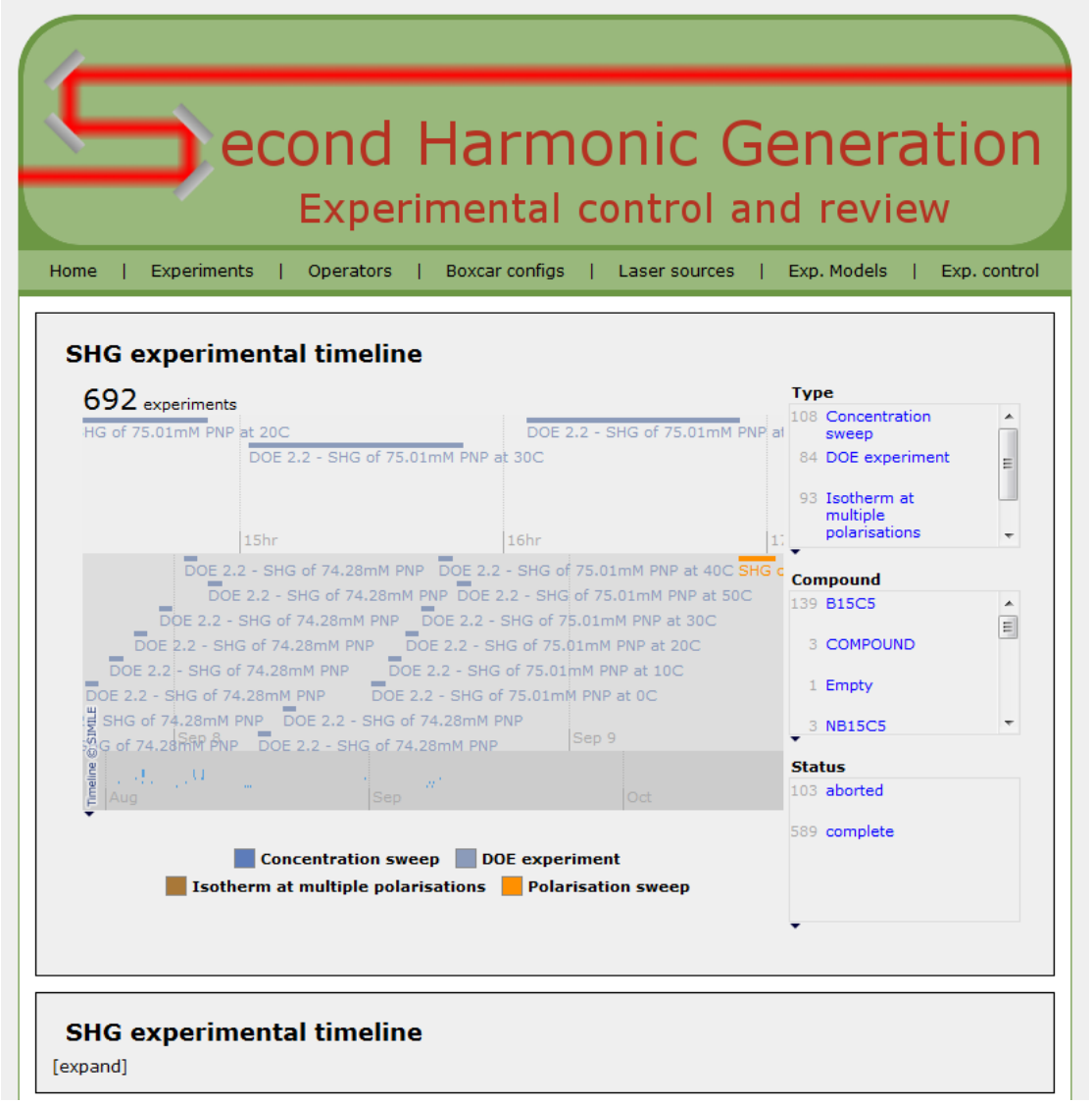


FIGURE 5.3: The SHG experimental time line

5.2.2 Faceted searching

In addition to the experimental time line, an option to use faceted searching of experiments was provided. This was achieved using the Simile Exhibit[87] software. As with the Timeline software, Exhibit loads a JSON representation of the experiments and uses this to generate its view. The developer specifies the facets to filter on as HTML DIVs, referring to elements in the JSON specification. The facets chosen for this work were experiment type, compound investigated and the status of the experiment. A number of views are available through the software including lists, interactive maps and time

lines. In this work the time line was used, as discussed above. Fig. 5.4 shows the time line after filtering on a number of facets.

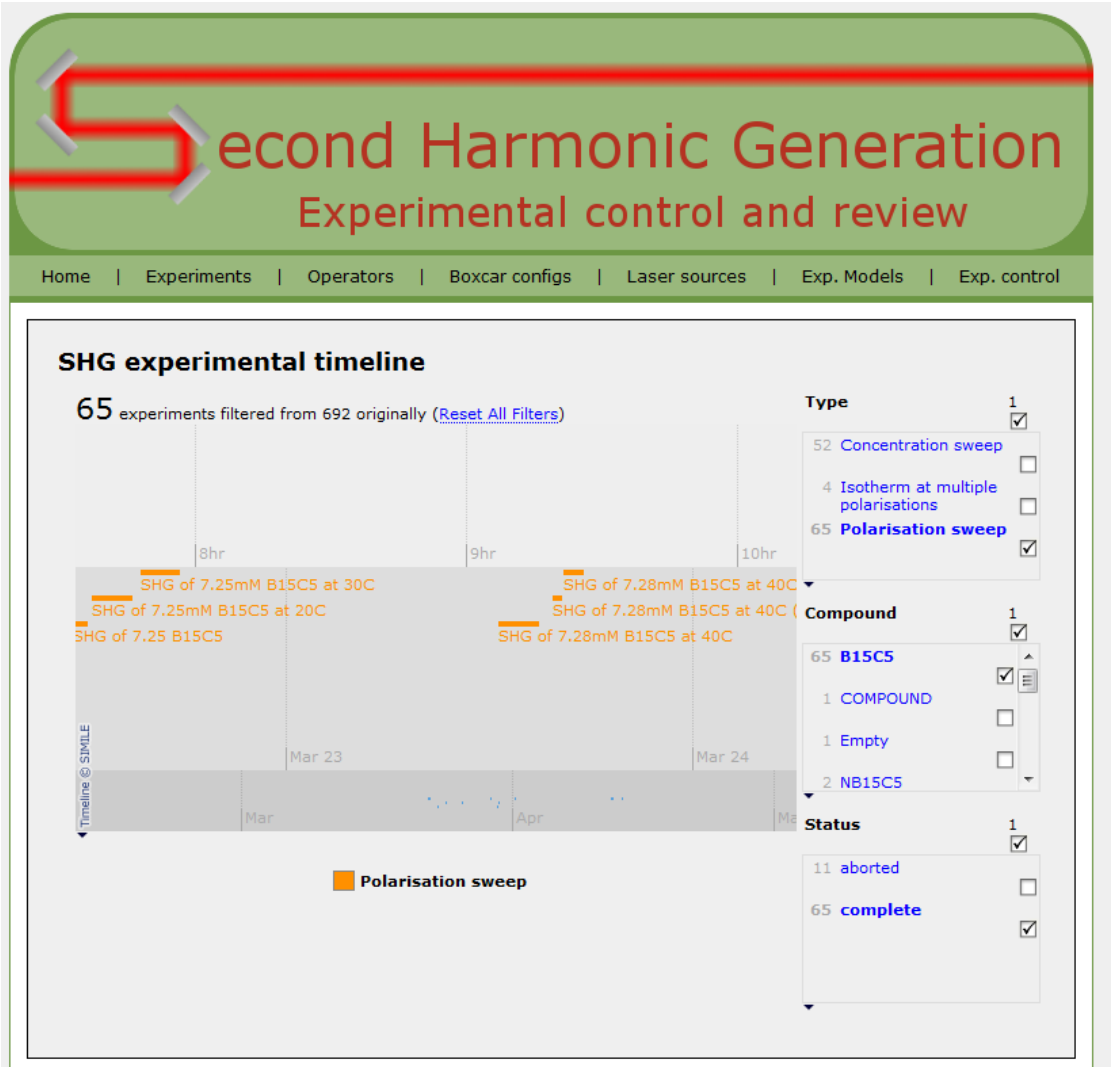


FIGURE 5.4: The SHG experimental time line after applying a number of filters via the faceted searching tool

5.2.3 Result plots

A plot of experimental results is shown in a number of locations on the website. The home page shows a 3 x 3 matrix of plots showing the nine most recent experiments, when selecting an experiment via the timeline a mini plot of the results in given and the experimental results pages shows a full plot of the data. These plots are generated on-the-fly directly from the results in the database, which allows real-time monitoring of the experimental results. The plots are generated using the Google Chart Tools[88], an API providing a wide range of tools to produce various charts type and diagrams. Each chart is generated via a number of parameters encoded into a single URL and submitted through a HTTP request. The chart is returned as a PNG image. Google have applied

a maximum of 2048 characters for a given URL, which limits the total amount of data available to plot in a single chart. When generated the URL for the run data (each laser shot of a given run) this limit was exceeded; to overcome this the larger datasets were filtered to show every n'th point, n depending on the size of the data. The plots shown on the experimental results page also provide mouse over functionality, such that when the mouse is over a data point a pop-up information box displays the details of that point. This mouse over functionality also provides links over the chart allowing a user to select a point on the graph and mine deeper into the data. This is generated through JavaScript, creating a map of points over the plot and applying mouse over functions.

5.2.4 Environmental data

The environmental data can be displayed in one of two ways depending on the state of the experiment. If the experiment is active the environmental data within the laboratory is displayed in tabular format. This is updated automatically every fifteen seconds providing a real-time monitor of conditions in the laboratory. This allows users to detect anomalies in the environmental conditions during the experiments. If an experiment has completed, a Simile TimePlot is generated showing a plot of all captured environmental conditions during the time frame of that experiment. This provides detailed information in changes of environmental conditions during a given experiment. This information can be used to explain unexpected results due to factors such as a large change in room temperature which will effect the temperature of the sample. Fig. 5.5 shows an example of the two displays of environmental conditions.

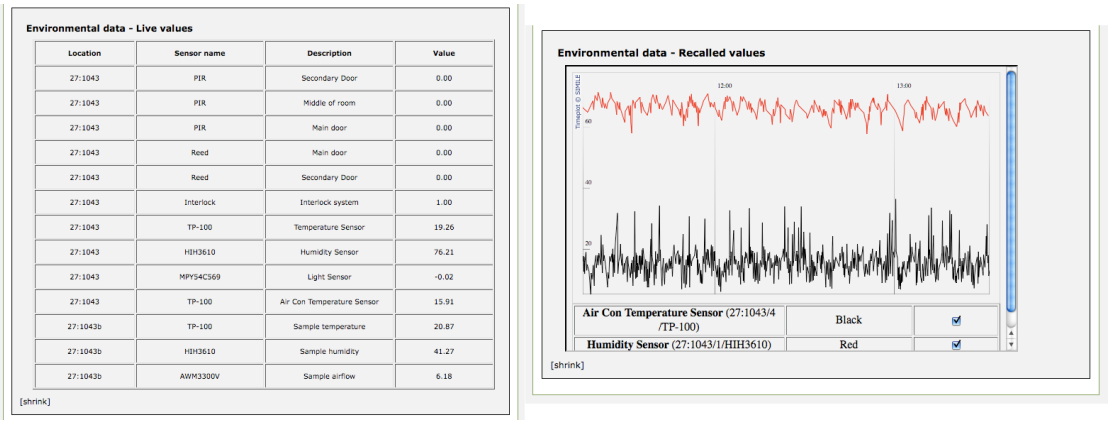


FIGURE 5.5: The SHG environmental conditions display within the web site for the period of a given experiment. Left: Active experiment tabular view. Right: Completed experiment Simile TimePlot

5.2.5 Data modelling

Once an experiment has completed the data is fitted via a non-linear regression to a model equation using a statistics software package, R[89]. This is discussed further in

chapter 6. This process was automated as part of the experimental workflow through use of the message broker. Once an experiment has completed a message is published to a 'modelling' script. This downloads the experimental data through the web interface and selects a model depending on the experiment type (also retrieved through the web interface). This is then submitted to the modelling software. When the modelling has complete the resulting plots and output is inserted into the SHG experimental database and are presented through the web interface. The available modelling results are displayed using the Simile Timeline software, as with the experimental results. The use of the message broker to make request to fit the data allows the modelling script to run on a dedicated server, eliminating the chance of slowing the web server. This process could be extended to making requests to the Grid with minimal adaptation.

Fig. 5.6 displays a number of screen shots from the SHG website, showing how each part of the system links together.

5.3 Discussion

The SHG website has proved to be a valuable resource for collecting and displaying the experimental information and its associated metadata. Providing a central point containing all this information allows for quick access to all the required data, either in real-time or when reviewing older experiments. The ability to monitor an experiment in real-time allows early detection of failures, preventing time from being wasted. Should a result be erroneous, the experimenter can review the environmental conditions and other metadata to determine the cause. Each page containing experimental results also provides a CSV download of that data via a unique URL. This allows external software to request the data in a known format, increasing the interoperability of the system.

When linked with an ELN, the website becomes the single link describing both the experimental process and results. The experimental link can be embedded into the ELN, where the discussion of this data is then recorded. This is extended further through linking the data modelling as part of the experimental process. The use of open source software to present the data has allowed for rapid development while maximising the functionality.

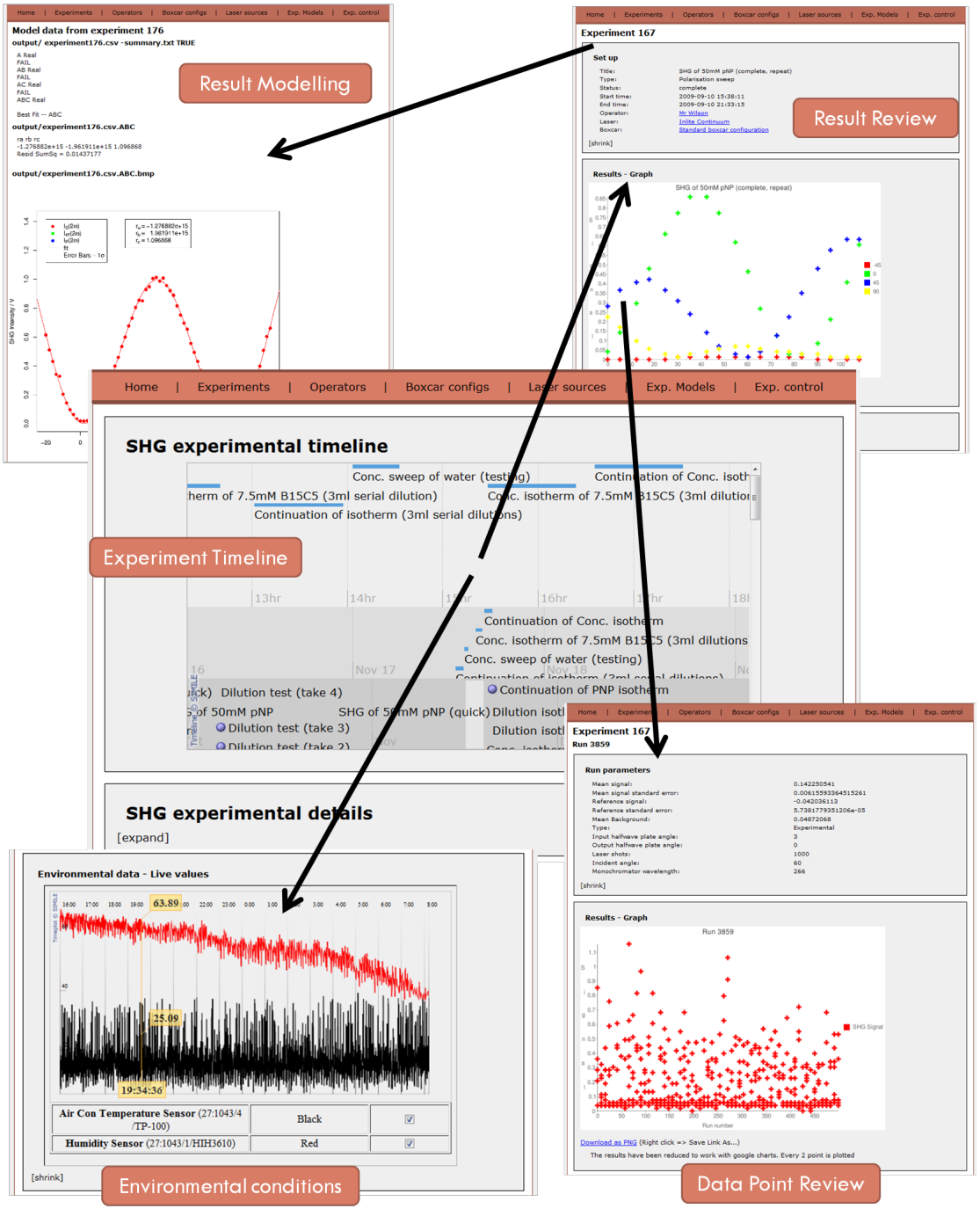


FIGURE 5.6: A number of screen shots showing how the components of the SHG website link together

Chapter 6

Second Harmonic Generation

6.1 Background

Non-linear optics is the study of phenomena that occur due to changes in the optical properties of a medium due to the presence of light. To observe these non-linear effects the light must be of a high enough intensity such as that of the amplitude supplied by a laser. The first reported non-linear effect was the phenomenon of second harmonic generation by Franken *et al.* in 1961[90]. Second harmonic generation is the conversion of two photons of frequency ω to a single photon of frequency 2ω. This is shown in Fig. 6.1.

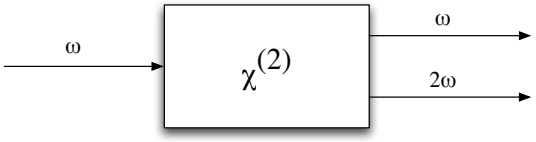


FIGURE 6.1: Representation of second harmonic generation through a medium

6.1.1 Theory

When light is incident on a dielectric medium the electric field induces a dipole moment through a change in the charge distribution. This dipole moment re-radiates the energy at the same frequency as the incident light. The total dipole moment per unit volume is called the electric polarisation, P, and is proportional to the incident electric field, E:

$$P = \epsilon_0 \chi^{(1)} E \tag{6.1}$$

where ϵ_0 is the permittivity of free space and $\chi^{(1)}$ is the electric susceptibility of the medium. With high intensity light, and therefore higher electric fields, non-linear effects are observed. The electric polarisation is then described as a power series:

$$P = \epsilon_0(\chi^{(1)}E + \chi^{(2)}E^2 + \chi^{(3)}E^3 + ...) \tag{6.2}$$

where $\chi^{(2)}$, $\chi^{(3)}$, are the second- and third-order non-linear optical susceptibilities of the medium. Considering the incident light as an electromagnetic wave, this can be written as:

$$E = E_0 \cos(\omega t) \tag{6.3}$$

where ω is the frequency of the incident light, t is the time and E_0 is the electric field at $t = 0$. By combining Eq. 6.2 and 6.3 the electric polarisation P can be described as:

$$P = \epsilon_0(\chi^{(1)}E_0 \cos(\omega t) + \chi^{(2)}E_0^2 \cos^2(\omega t) + \chi^{(3)}E_0^3 \cos^3(\omega t) + ...) \tag{6.4}$$

The second harmonic generation effect is governed by the second order polarisation term:

$$P^{(2)} = \epsilon_0\chi^{(2)}E_0^2 \cos^2(\omega t) \tag{6.5}$$

which can be re-written as:

$$P^{(2)} = \frac{1}{2}\epsilon_0\chi^{(2)}E_0^2[\cos(2\omega t) + 1] \tag{6.6}$$

This equation shows that there is a relationship between the electric polarisation and light with twice the frequency of the incident light. It can also be seen that the intensity of the effect is dependent on the square of the intensity of the incident light, hence a non-linear effect.

As shown in Eq. 6.1 an applied field, *E*, will result in an electric polarisation, *P*. If applying the electric field in the opposite direction, -*E*, the electric polarisation will also be inverted, -*P*, in an isotropic medium. This also holds true for the second harmonic polarisation term, $P^{(2)}$. By following through the calculation with *E* and -*E* in Eq. 6.5, it is shown that $P^{(2)}$ must equal -$P^{(2)}$. This can only hold true when $\chi^{(2)}$ is zero. This shows that media with a centre of inversion, such as bulk liquid or gases, do not produce second harmonics or high order even power non-linear effects. However, at the interface between two isotropic regions the symmetry is lost and is therefore anisotropic. This

makes second harmonics generation an interface specific technique, and in this work will be used to investigate the liquid-air interface.

6.1.2 Molecular orientation at the surface

The intensity of the generated second harmonic at an interface can be determined using Eq. 6.7, derived by Mizrahi and Sipe in 1988[91].

$$I_{2\omega} = \frac{32\pi^3\omega^2}{c^3} \sec^2\theta_{in} |\hat{e}_{2\omega} \cdot \chi^{(2)} : \hat{e}_w \hat{e}_w|^2 I_\omega^2 \tag{6.7}$$

where the subscripts ω and 2ω are values at the fundamental and second harmonic wavelength respectively, θ is the angle of the beam to the surface normal of the sample, c is the speed of light, $\chi^{(2)}$ is the second-order non-linear optical susceptibilities of the medium and $\hat{e}$ relates to the polarisation of the incident beam and the Fresnel coefficients as described in Eq. 6.8.

$$\hat{e} = \bar{e}.e \tag{6.8}$$

were $\bar{e}$ is the unit vector of the polarisation of the incident light and e is the corresponding Fresnel coefficients.

The vector $\hat{e}$ can be described in reference to a Cartesian coordinate system (x,y,z), whereas light is traditionally described in terms of s- and p-polarisation. To link these terms the orientation must be described. These are traditional described with the Z-axis as the surface normal and the XZ plane parallel to the plane of incidence of the beam. A P-polarised beam will oscillate in the XZ plane (perpendicular to the surface of the sample) and a S-polarised beam will oscillate parallel to the Y-axis (in the plane of the surface of the sample). This is shown in Fig. 6.2. γ will be used to describe the rotation away from P-polarisation of the incident beam where P-polarisation is 0^o and S-polarisation is 90^o. Using this definition we can write ê as a vector:

$$\hat{e} = \begin{bmatrix} E_{X\omega} \\ E_{Y\omega} \\ E_{Z\omega} \end{bmatrix} = \begin{bmatrix} \cos(\gamma) e_{X\omega} \\ -\sin(\gamma) e_{Y\omega} \\ \cos(\gamma) e_{Z\omega} \end{bmatrix} \tag{6.9}$$

In this three-dimensional system P and E have direction associated with them, as shown above, and are represented as vectors. χ is now linking one vector to the product of two others, $|\hat{e}_{2\omega} \cdot \chi^{(2)} : \hat{e}_w \hat{e}_w|$, and becomes a third-rank tensor. Eq. 6.5 can now be written in the general form:

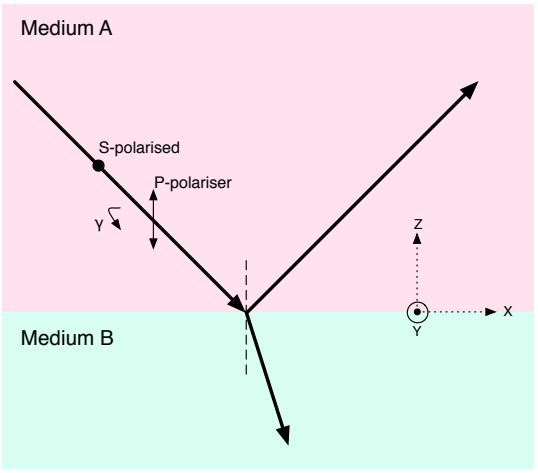


FIGURE 6.2: Orientation and polarisations of the incident beam to the sample

$$P^{(2)} = \sum_{jk} \chi^{(2)}_{ijk} E_j(\omega) E_k(\omega) \tag{6.10}$$

where i, j and k have values of X, Y, and Z (with orientation as described above). The χ_{ijk} tensor can be reduced to four elements when it is assumed the surface is invariant with rotation about the surface normal[92]. This is reduced further when the system is achiral, a mirror plane is present, as the χ_{XYZ} component becomes zero[93]. With this assumption the polarisation, $P^{(2)}$, can be calculated as:

$$\begin{bmatrix} P_X \\ P_Y \\ P_Z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \chi_{XYZ} & \chi_{XXZ} & 0 \\ 0 & 0 & 0 & \chi_{XXZ} & -\chi_{XYZ} & 0 \\ \chi_{XXZ} & \chi_{ZXX} & \chi_{ZZZ} & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} E^2_{X\omega} \\ E^2_{Y\omega} \\ E^2_{Z\omega} \\ 2E_{Y\omega}E_{Z\omega} \\ 2E_{X\omega}E_{Z\omega} \\ 2E_{X\omega}E_{Y\omega} \end{bmatrix} \tag{6.11}$$

Using this component form it follows that when a sample is incident with a P-polarised beam ($E_x = E_y = 0$) then only a P-polarised component will be observed ($P_x = P_y = 0$ and $P_z = \chi_{ZZZ}$). Similarly when applying a S-polarised beam ($E_x = E_z = 0$) only a P-polarised beam is observed ($P_x = P_y = 0$ and $P_z = \chi_{ZXX}$). To observe an output of S-polarisation a beam with both S- and P-polarisation components must be used. Using this theory and combining with Eq. 6.9 and 6.7 it can be shown that:

$$I_p(2\omega) \propto \left| A\cos^2\gamma + B\sin^2\gamma \right|^2 I^2(\omega) \tag{6.12}$$

$$I_s(2\omega) \propto |C \sin 2\gamma|^2 \, I^2(\omega) \tag{6.13}$$

$$I_{+45}(2\omega) \propto \left| \frac{1}{\sqrt{2}} (A \cos^2 \gamma + B \sin^2 \gamma + C \sin 2\gamma) \right|^2 I^2(\omega) \tag{6.14}$$

$$I_{-45}(2\omega) \propto \left| \frac{1}{\sqrt{2}} (A \cos^2 \gamma + B \sin^2 \gamma - C \sin 2\gamma) \right|^2 I^2(\omega) \tag{6.15}$$

$$A = a_2 \chi_{XXZ} + a_3 \chi_{ZXX} + a_4 \chi_{ZZZ} \tag{6.16}$$

$$B = a_5 \chi_{ZXX} \tag{6.17}$$

$$C = a_1 \chi_{XXZ} \tag{6.18}$$

where a_1 - a_5 are functions of the refractive indices and angle of incidence and represent the electric field at the surface. These are calculated by[94]:

$$\begin{aligned} a_1 &= e_y(\omega) e_z(\omega) e_y(2\omega) \\ a_2 &= 2 e_x(\omega) e_z(\omega) e_x(2\omega) \\ a_3 &= |e_x(\omega)|^2 e_z(2\omega) \\ a_4 &= |e_z(\omega)|^2 e_z(2\omega) \\ a_5 &= |e_y(\omega)|^2 e_z(2\omega) \\ a_6 &= -2 e_y(\omega) e_z(\omega) e_y(2\omega) \\ a_7 &= e_y(\omega) e_z(\omega) e_x(2\omega) \end{aligned} \tag{6.19}$$

where:

$$e_x(\omega) = \cos(\theta_{in})\{1 - \frac{n_2(\omega)\cos(\theta_{in}) - n_1(\omega)\cos(\theta_{ref})}{n_1(\omega)\cos(\theta_{ref}) + n_2(\omega)\cos(\theta_{in})}\}$$

$$e_x(2\omega) = -\frac{2n_1(2\omega)\cos(\theta_{in})\cos(\theta_{out})}{n_1(2\omega)\cos(\theta_{ref}) + n_2(2\omega)\cos(\theta_{in})}$$

$$e_y(\omega) = 1 + \frac{n_1(\omega)\cos(\theta_{in}) - n_2(\omega)\cos(\theta_{ref})}{n_1(\omega)\cos(\theta_{in}) + n_2(\omega)\cos(\theta_{ref})}$$

$$e_y(2\omega) = \frac{2n_1(2\omega)\cos(\theta_{in})}{n_1(2\omega)\cos(\theta_{in}) + n_2(2\omega)\cos(\theta_{ref})}$$

$$e_z(\omega) = \frac{n_1(\omega)^2\sin(\theta_{in})}{n_3(\omega)^2}\{1 + \frac{n_2(\omega)\cos(\theta_{in}) - n_1(\omega)\cos(\theta_{ref})}{n_1(\omega)\cos(\theta_{ref}) + n_2(\omega)\cos(\theta_{in})}\}$$

$$e_z(2\omega) = \frac{1}{n_3(2\omega)^2}\{\frac{2n_1(2\omega)\cos(\theta_{in})\sin(\theta_{out})}{n_1(2\omega)\cos(\theta_{ref}) + n_2(2\omega)\cos(\theta_{in})}\} \tag{6.20}$$

where θ_{in} in the angle of incidence, θ_{out} is the angle of reflection, θ_{ref} is the angle of refraction and $n_i(\omega)$ and $n_i(2\omega)$ are the refractive indices at fundamental and second harmonic frequencies of the interfacial system shown in Fig. 6.3.

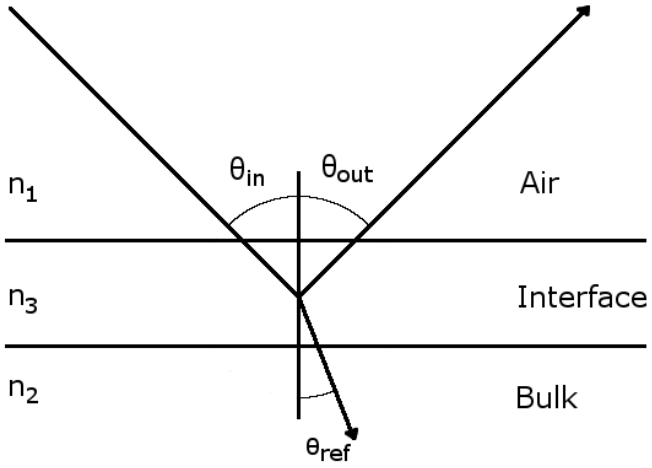


FIGURE 6.3: Illustration of the three layer model assumed in the SHG experiments

By combining eq. 6.12 and 6.13 we can determine a single equation for an arbitrary output polarisation, Γ. We must first define a point of reference, in this case at the detector with P-polarised second harmonic signal at zero Γ[94]. A signal at arbitrary Γ will be contain a component of both S- and P-polarisation. This is shown in Fig. 6.4. The combined fitting equation is then given by:

$$I(2\omega) = |(C\sin 2\gamma)\sin\Gamma + (A\cos^2\gamma + B\sin^2\gamma)\cos\Gamma|^2 \tag{6.21}$$

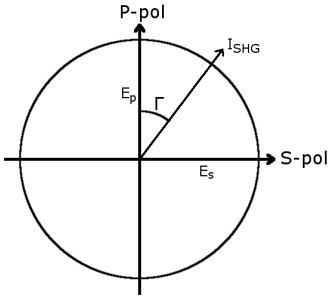


FIGURE 6.4: A plot to show the angular dependence of the second harmonic signal to the output polarisation, Γ, with respect to the S- and P-polarised components

The coefficients A, B and C can be determined experimentally through detection of the second harmonic at a range of input polarisations between S and P at a number of output polarisation (S, P, $+45^o$ and -45^o) and fitting this data to Eq. 6.21. This experimental practise has been optimised using DOE methodologies, discussed in chapter 11. When resonance enhancement is present these A, B and C values become complex. From these fitted values the second-order non-linear optical susceptibilities of the medium, $\chi^{(2)}$, can be determined by rearranging Eq. 6.16, 6.17 and 6.18.

Once the second-order non-linear optical susceptibilities of the medium, $\chi^{(2)}$, have been experimentally determined it can be related to the hyperpolarisability, β_{xyz} of the molecule. This can be used to estimate the average orientation of the molecules at the surface. The relationship between $\chi^{(2)}$ and β_{xyz} given by:

$$\chi^{(2)}_{XYZ} = N_s \sum_{xyz} \langle \mathbf{T}^{xyz}_{XYZ}(\phi, \theta, \psi) \rangle \beta_{xyz} \tag{6.22}$$

where N_s is the surface density, $\mathbf{T}$ is a transformation matrix between the laboratory orientation, (X, Y, Z) and the surface molecular orientation (x, y, z) and the '$\langle\rangle$' indicate averaging over the molecular orientation distribution. θ is the angle between Z and z, ψ is the rotation about the Z-axis and ϕ is the azimuthal angle. This is shown in Fig. 6.5.

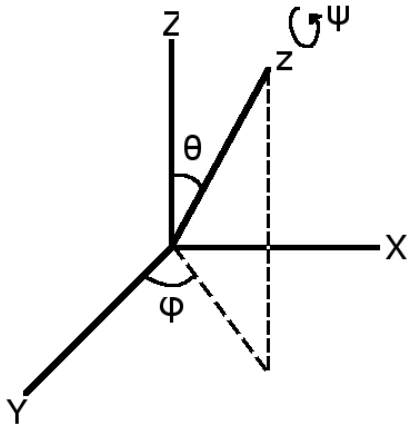


FIGURE 6.5: Illustration of the angles relating the molecular orientation to the laboratory frame

The susceptibility terms, χ_{zxx}, χ_{xxz} and χ_{zzz}, can be expressed in terms of the rotational parameters between the laboratory and molecular orientation, θ, ψ and ϕ and the hyperpolarisability terms, β_{zxx}, β_{xxz}, β_{xzx} and β_{zzz}[94]. At the liquid/air interface it can be considered that the surface is invariant with rotation about the the surface normal, therefore the equations can be expressed independent of the azimuthal angle, ϕ. These equations are shown in Eq. 6.23, 6.24 and 6.25.

$$\chi_{xxz} = \frac{N_s}{2\epsilon_0}[\langle\sin^2(\theta)\cos(\theta)\rangle\beta_{zzz} - \langle\sin^2(\psi)\cos(\theta)\sin^2(\theta)\rangle(\beta_{zxx} + 2\beta_{xzx}) + \langle\cos(\theta)\rangle\beta_{xxz}] \quad (6.23)$$

$$\chi_{zxx} = \frac{N_s}{2\epsilon_0}[\langle\sin^2(\theta)\cos(\theta)\rangle\beta_{zzz} - \langle\sin^2(\psi)\cos(\theta)\sin^2(\theta)\rangle(\beta_{zxx} + 2\beta_{xzx}) + \langle\cos(\theta)\rangle\beta_{zxx}] \quad (6.24)$$

$$\chi_{zzz} = \frac{N_s}{\epsilon_0}[\langle\cos^3(\theta)\rangle\beta_{zzz} - \langle\sin^2(\psi)\cos(\theta)\sin^2(\theta)\rangle(\beta_{zxx} + 2\beta_{xzx}) \quad (6.25)$$

where N_s is the number of active molecules per unit surface and ϵ_0 is the permittivity.

The angle between the surface normal, Z, and the main molecular axis, z, θ, will be the focus of the research completed on surface orientation. This parameter describes how flat, on average, the molecule is at the interface. The surface orientation parameters, D, is calculated from Eq. 6.26 and describes a Dirac distribution of this angle.

$$D = \frac{\langle \cos^3(\theta) \rangle}{\langle \cos(\theta) \rangle} \tag{6.26}$$

Through a number of assumptions, discussed below, the experimentally determined molecular susceptibility, χ_{xyz}, can be related to the hyperpolarisability, β_{xyz}, and molecular orientation. The general form for a rotationally isotropic distribution of molecules at the interface is shown in Eq. 6.27 as the ratio between the susceptibility terms[95].

$$\begin{aligned} \frac{\chi_{zxx}}{\chi_{xxz}} &= \frac{(\beta_1 + \beta_2 - 2\beta_3)D - (\beta_1 - \beta_2 - 2\beta_3)}{(\beta_1 - \beta_2)D - (\beta_1 - \beta_2 - 2\beta_3)} \\ \frac{\chi_{zzz}}{\chi_{xxz}} &= \frac{2(\beta_2 + 2\beta_3)D + 2(\beta_1 - \beta_2 - 2\beta_3)}{(\beta_1 - \beta_2)D - (\beta_1 - \beta_2 - 2\beta_3)} \end{aligned} \tag{6.27}$$

where β_1, β_2 and β_3 are determined by the symmetry of the molecule under investigation. Zhang *et al.*[95] detail the independent non-zero hyperpolarisability tensor elements for a range of Cn symmetries and how they relate to Eq. 6.27. This work will focus on molecules with C2v and therefore any further discussion will assume C2v symmetry.

An assumption about the distribution of the angle describing the rotation about the main molecular angle, ψ, must also be made to relate the molecular susceptibility, χ_{xyz}, to the hyperpolarisability, β_{xyz}. In general three cases are considered; a delta Dirac distribution of ψ centred on values 90^o and 0^o or a random distribution. Table 6.1 shows the transformation from hyperpolarisability terms to rotation angle ψ for Eqs. 6.23, 6.24 and 6.25 for these three cases.

χ_{xyz}	ψ	β_{xzx}	β_{zxx}	β_{zzz}
xxz	0^o	$\frac{1}{2}\cos(\theta)$	0	$\frac{1}{2}\cos(\theta)\sin^2(\theta)$
	90^o	$\cos^3(\theta) - \frac{1}{2}\cos(\theta)$	$-\frac{1}{2}\cos(\theta)\sin^2(\theta)$	$\frac{1}{2}\cos(\theta)\sin^2(\theta)$
	random	$\frac{1}{2}\cos^3(\theta)$	$-\frac{1}{4}\cos(\theta)\sin^2(\theta)$	$\frac{1}{2}\cos(\theta)\sin^2(\theta)$
zxx	0^o	0	$\frac{1}{2}\cos(\theta)$	$\frac{1}{2}\cos(\theta)\sin^2(\theta)$
	90^o	$-\cos(\theta)\sin^2(\theta)$	$\frac{1}{2}\cos^3(\theta)$	$\frac{1}{2}\cos(\theta)\sin^2(\theta)$
	random	$-\frac{1}{2}\cos(\theta)\sin^2(\theta)$	$\frac{1}{4}(\cos(\theta) + \cos^3(\theta))$	$\frac{1}{2}\cos(\theta)\sin^2(\theta)$
zzz	0^o	0	0	$\cos^3(\theta)$
	90^o	$2\cos(\theta)\sin^2(\theta)$	$\cos(\theta)\sin^2(\theta)$	$\cos^3(\theta)$
	random	$\frac{1}{2}\cos(\theta)\sin^2(\theta)$	$\cos(\theta)\sin^2(\theta)$	$\cos^3(\theta)$

TABLE 6.1: The transformation matrix of the susceptibility to the hyperpolarisability for the three assumptions of ψ

When considering a molecule with C_{2V} symmetry that is resonant at the second harmonic wavelength, Eq. 6.28 is used to calculate the ratio between the dominant hyperpolarisability terms, β_{zxx} and β_{zzz}. The molecular tilt angle is determined with Eq.

6.29 when considering ψ, the rotation of the molecule about its Z-axis, to be random and Eq. 6.30 when taking this angle to be fixed at 90^o .

$$\frac{\beta_{zxx}}{\beta_{zzz}} = 2\frac{\chi_{zxx} - \chi_{xxz}}{\chi_{zzz} + 2\chi_{xxz}} \tag{6.28}$$

$$D = \frac{\chi_{zzz} - 2\chi_{zxx} + 2\chi_{xxz}}{\chi_{zzz} - 2\chi_{zxx} + 4\chi_{xxz}} \tag{6.29}$$

$$D = \frac{\chi_{zzz} - \chi_{zxx} + \chi_{xxz}}{\chi_{zzz} - \chi_{zxx} + 3\chi_{xxz}} \tag{6.30}$$

Through making the assumptions discussed above about ψ and which hyperpolarisability terms, β_{ijk}, are dominant due to the molecular symmetry, Eqs. 6.23, 6.24 and 6.25 can be expressed in terms of F, R and D[96] where:

$$F = \frac{N_s}{2\epsilon_0}\langle\cos(\theta)\rangle\beta_{zzz} \tag{6.31}$$

$$R = \frac{\beta_{zxx}}{\beta_{zzz}} \tag{6.32}$$

$$D = \frac{\langle\cos^3(\theta)\rangle}{\langle\cos(\theta)\rangle} \tag{6.33}$$

For the assumption that β_{zzz} and β_{zxx} will be dominant:

In the case of ψ considered random, such that $\langle\sin^2(\psi)\rangle = 0.5$, Eq. 6.23, 6.24 and 6.25 can be expressed in terms of F, D and R as:

$$\chi_{xxz} = F[(1-D)(1-\frac{R}{2})] \tag{6.34}$$

$$\chi_{zxx} = F[(1-D)(1-\frac{R}{2}) + R] \tag{6.35}$$

$$\chi_{zzz} = F[2D + (1-D)R] \tag{6.36}$$

Similarly when ψ is fixed to 90^o, and $\langle\sin^2(\psi)\rangle = 1$, Eq. 6.23, 6.24 and 6.25 can be expressed in terms of F, D and R as:

$$\chi_{xxz} = F[(1-D)(1-R)] \tag{6.37}$$

$$\chi_{zxx} = F[(1-D)(1-R)+R] \tag{6.38}$$

$$\chi_{zzz} = F[2D+(1-D)2R] \tag{6.39}$$

Through combining these equations with Eq. 6.21, 6.16, 6.17 and 6.18 an equation can be described to fit the experiments data directly to the orientation and hyperpolarisability parameters, in the case of random ψ:

$$\begin{aligned} I(2\omega) = |a_1 F(1-D)(1-\frac{R}{2})\sin(2\gamma)\sin(\Gamma) + ((a_2 F(1-D)(1-\frac{R}{2}) \\ +a_3 F((1-D)(1-\frac{R}{2})+R) + a_4 F(2D+(1-D)R)) * \cos(\gamma)^2 \\ +a_5 F((1-D)(1-\frac{R}{2})+R) * \sin(\gamma)^2) * \cos(\Gamma)|^2 \end{aligned} \tag{6.40}$$

6.1.3 Adsorption at the interface

SHG can be used to study the equilibrium of molecules adsorbing at the interface. This is conducted by monitoring second harmonic signal intensity as a function of bulk concentration. The Langmuir adsorption model[97] assumes the surface contains a finite number of equivalent sites for the molecules to adsorb into as a monolayer. The observed second harmonic intensity is proportional to the square of the number of molecules adsorbed at the interface:

$$I_{SHG} = s\Theta^2 \tag{6.41}$$

where Θ is the fractional coverage of the interface and s is a scaling factor. The adsorption of a molecule is treated as a reversible reaction where a molecule, M, associates to an empty site, S_{emtpy} to produce a filled site, S_{filled}:

$$M + S_{empty} \rightleftharpoons S_{filled} \tag{6.42}$$

and the rates of these processes given by:

$$R_{ads} = k_{ads}[M](1-\Theta) \tag{6.43}$$

$$R_{des} = k_{des}\Theta \tag{6.44}$$

At equilibrium these rates are equal, allowing the equations given in Eq. 6.43 and 6.44 to be combined to:

$$\Theta = \frac{k[M]}{1 + k[M]} \tag{6.45}$$

where k is the adsorption equilibrium constant ($\frac{k_{ads}}{k_{des}}$). This is the Langmuir equation. As the concentration of the sample is lowered, the coverage at the surface will be lower and as such the signal will decrease. As there will also be solvent molecules at the surface, a proportion of the second harmonic will be generated from these. At the lower sample concentrations the second harmonic signal will be predominately from these solvent molecules. This is shown by including the susceptibilities of the solvent in a simplified version of Eq. 6.7.

$$I_{SHG} \propto |\chi_{Solvent} + \chi_{Sample} \cdot \Theta|^2 \tag{6.46}$$

Through multiplying out Eq. 6.46 and combining with Eq. 6.45 the second harmonic intensity can be described as:

$$I_{SHG} = A^2 + \frac{2A\cos(\phi)Sk[M]}{1 + k[M]} + (\frac{Sk[M]}{1 + k[M]})^2 \tag{6.47}$$

where A is the susceptibility of the solvent, S is the susceptibility of a monolayer of the sample and ϕ is the phase difference between sample and solvent. In this equation the scaling factor has been included in the susceptibilities as it cannot be determined experimentally.

6.2 Experimental apparatus

6.2.1 Optical configuration

The SHG experimental set up was as shown in Fig. 6.6.

The laser source used in the experiments was a Continuum Inlite; a Q-switched Nd:YAG pulsed laser which is internally doubled to produce a beam at 532 nm. The laser has a pulse width of 6 to 8 ns with energy of approximately 130 mJ per pulse and at a pulse rate of 20 Hz.

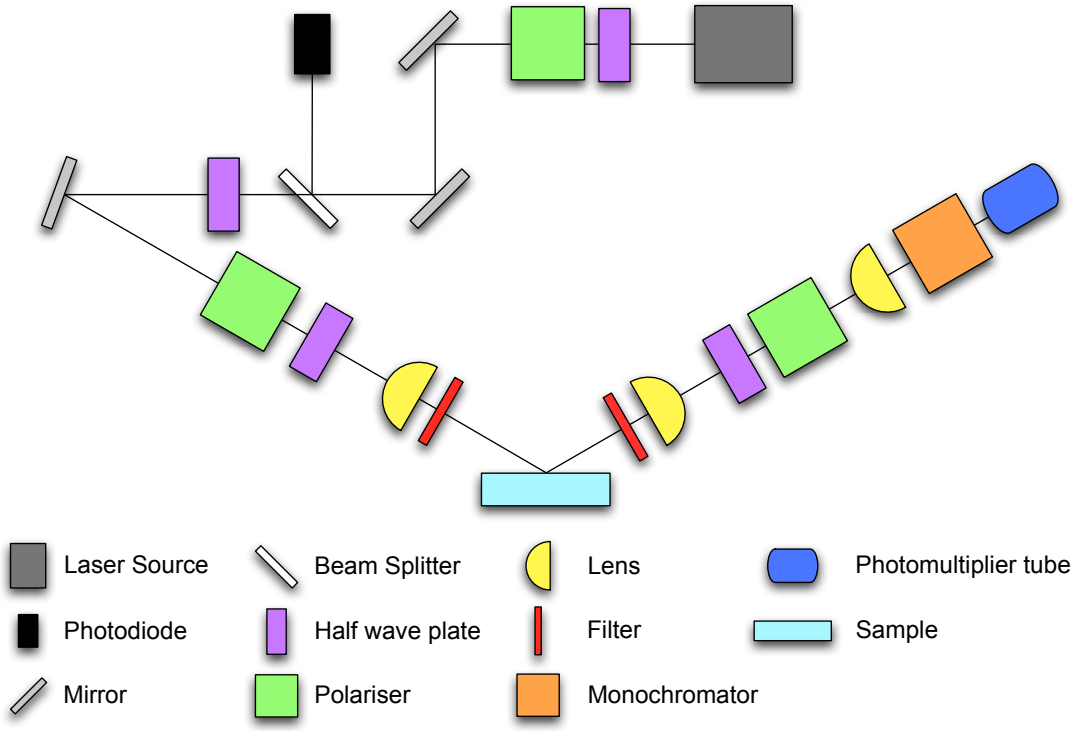


FIGURE 6.6: The optical set up used for the SHG experiments

The beam first passes through a half wave plate and splitting polariser, which acts as a power modulator, bringing the power down to experimental levels. The beam then passes through a glass plate, diverting 8% of the power to a photodiode. This provided a reference to the raw power output of the laser during the experiments. The remaining 92% of the beam is directed using a number of mirrors to the experimental rig, which is a 'two arm' set up aligned at 60^o to the surface normal of the sample. The first arm controls the incident light with a second power modulator providing fine resolution of the power, a 532 nm half wave plate controlling the polarisation of the incident light, a focusing lens and a yellow glass Schott UV filter to remove any second harmonic generated by the optics. The second arm control the second harmonic output with a UG5 filter removing any reflected 532 nm light, a re-collimating lens, a 266 nm half wave plate to control the polarisation and a rochon polariser. The ordinary path from this polariser passes through a focusing lens, through a monochromator and is incident on a photomultiplier tube. This optical arrangement was based on that described by Lefteris Danos[98].

The sample is contained within a glass Petri dish and covered by a plastic lid to minimise evaporation. The portion of the beam that is transmitted through the sample is blocked using a beam dump.

6.2.2 Electronic configuration

The photodiode and photomultiplier tube were connected to a National Instruments DAQPad-6020E via a Standford Research Systems boxcar using a SR250 gated integrator module. The trigger of the data acquisition card and boxcar integrator were synchronised to the laser pulse via the Q-switch output of the laser controller. The electrical connectivity diagram can be seen in Fig. 6.7. The data acquisition card was connected to a personal computer via USB and the results collect via software described in chapter 5.

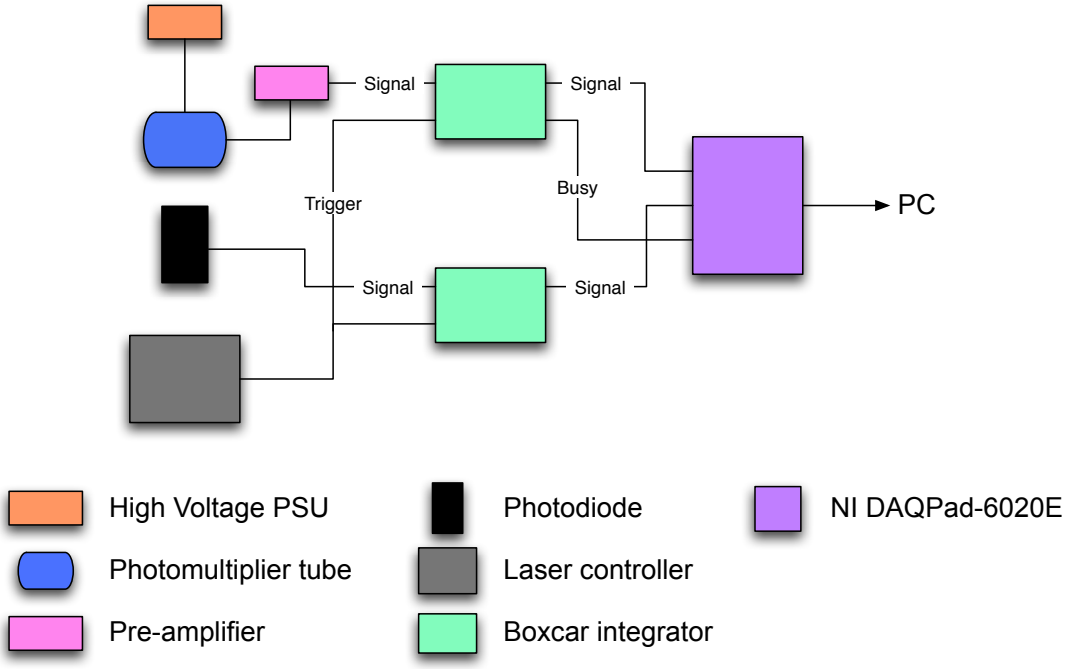


FIGURE 6.7: The electronic set up used for the SHG experiments

6.3 Alignment

6.3.1 Faraday modulator

6.3.1.1 Background

In has been shown that the use of a polarisation modulator combined with other optical components can provide a highly accurate determination of the polarisation of a beam of light[99, 100]. In this work a Faraday modulator was used to oscillate the polarisation of the laser beam while aligning the optical components. This method was used to align the polarisers and half wave plates to within 0.5^o of the true zero points. The Faraday modulator consists of a dielectric material (in this case a Terbium Gallium

Garnet crystal) and a magnetic field applied parallel to the propagation of the beam. As the beam propagates through the medium the polarisation is rotated; the magnitude of this can be calculated using Eq. 6.48.

$$\theta = BVl \tag{6.48}$$

where B is the magnitude of the magnetic field in Tesla, V is the Verdet constant for the material at the beam wavelength and l is the length of the medium within the magnetic field. The Faraday modulator oscillates the direction of magnetic field at a rate similar to that of the pulse rate of the laser. When placed between two polarisers the modulator can be used to determine if polarisers are crossed, aligned or somewhere in-between by observing the signal on an oscilloscope.

When considering the signal strength of a beam travelling through two polarisers it is expected that it will follow a sinusoidal wave with a maximum observed when the polarisers are parallel and a minimum when crossed. When applying the modulation to this beam the signal will oscillate either side of the expected value. When the polarisers are between parallel and crossed the scope will show the fundamental wave, as shown in Fig. 6.8.

As the polarisers approach either parallel or crossed the oscillations will cause the polarisation of the incident light to pass through the minimum/maximum position. This will be shown on the scope as two peaks of non-equal magnitude at twice the frequency of modulation. This is shown in Fig. 6.9.

As the polarisers approach the true aligned positions the peak magnitudes will converge. At the point of true alignment the peaks will be of equal magnitude, as the oscillations will move the signal equally either side of the minimum/maximum. This is shown in Fig. 6.10.

Using this system, it can be difficult to distinguish between when the polarisers are parallel or crossed. In the set up used in this work this was not an issue as the detector (a photodiode) would become saturated as approaching the parallel position and not show the frequency doubling effect. Therefore all polarisers were aligned at the crossed position.

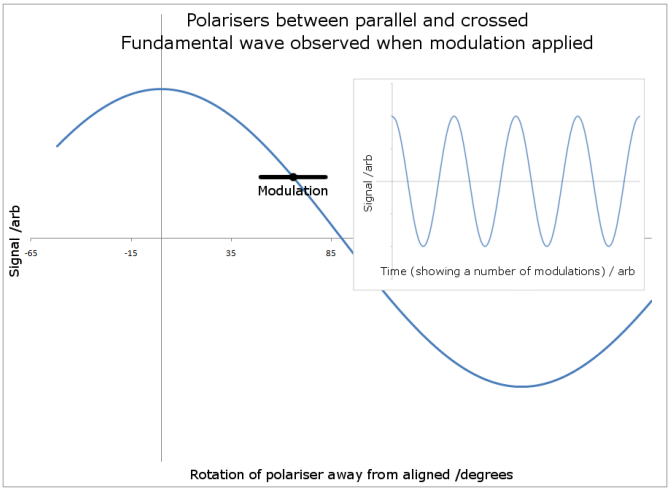


FIGURE 6.8: The fundamental wave observed when far from parallel/crossed polarisers when using a Faraday modulator

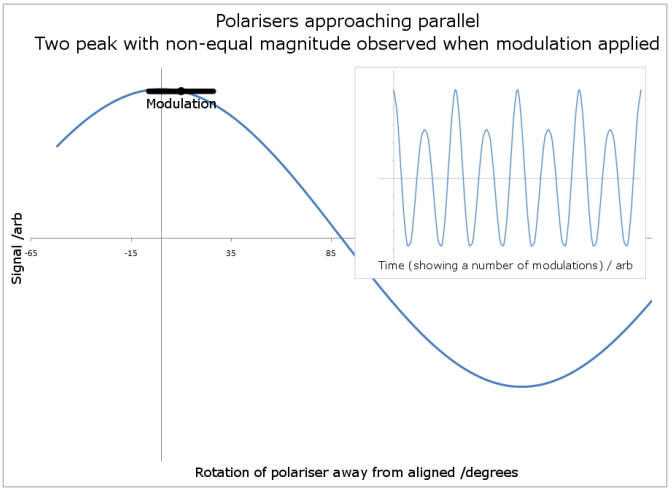


FIGURE 6.9: The frequency doubling effect with non-equal magnitude observed when modulating a beam close the parallel/crossed alignment

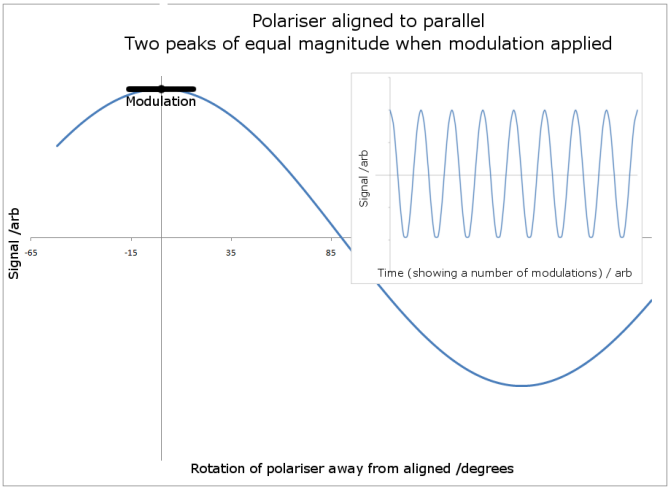


FIGURE 6.10: The frequency doubling effect with equal magnitude observed when modulating a beam at an aligned polariser position

6.3.1.2 Process

As the modulator requires two polarisers to show any effect, aligning the first optic is the most difficult procedure. This is achieved using a sample of water and setting the beam to the Brewster angle of the air/water interface. At this angle the water acts as a polariser, reflecting only s-polarised light. This allows the input polariser to be accurately aligned to p-polarisation by looking for the equal magnitude frequency doubling effect produced by the modulator. Once this has been achieved the beam path was modified to travel straight through both arms of the rig. This allowed for the output polariser to be aligned to s-polarisation by again rotating until the equal magnitude frequency doubling was observed. The zero point of both half wave plates was also determined using this method. The alignment was carried out with a 5 mW HeNe laser producing a beam at 632.8 nm. As the half wave plates are designed for 532 nm and 266 nm light their effect on the beam could not be quantified, therefore these were aligned by looking for the point where they had no effect on the polarisation of the beam. This was designated the zero point.

6.3.2 Fine adjustments

As shown in Eq. 6.13 when scanning the input polarisation at an output corresponding to S-polarisation, the signal should peak at 45^o and be zero at 0^o and 90^o. To determine the true alignment of the half wave plates a sample of PNP was prepared and a scan of input polarisation between -10^o and 100^o at S- output polarisation was completed. The signal intensity was plotted against the values generated by Eq. 6.49. If the system was aligned the plot would show a straight line for all points, whereas if incorrect the plot would show an ellipse.

$$\gamma' = \sin^2(2\gamma) \tag{6.49}$$

A plot of the results is shown in Fig. 6.11. It shows that the alignment is close but required a little tuning. To determine the required adjustment an offset was added to γ until a straight line was given, in this instance an offset of -0.5^o was required. The adjusted results are shown in Fig. 6.12. As shown in the plot the results between -10 to 0 and 90 to 100 do not align with the remaining results. It is believed this is due to the mounting of the half wave plate not sitting completely perpendicular to the direction of the beam. As this is not within the experimental range it has been assumed it will not effect the results.

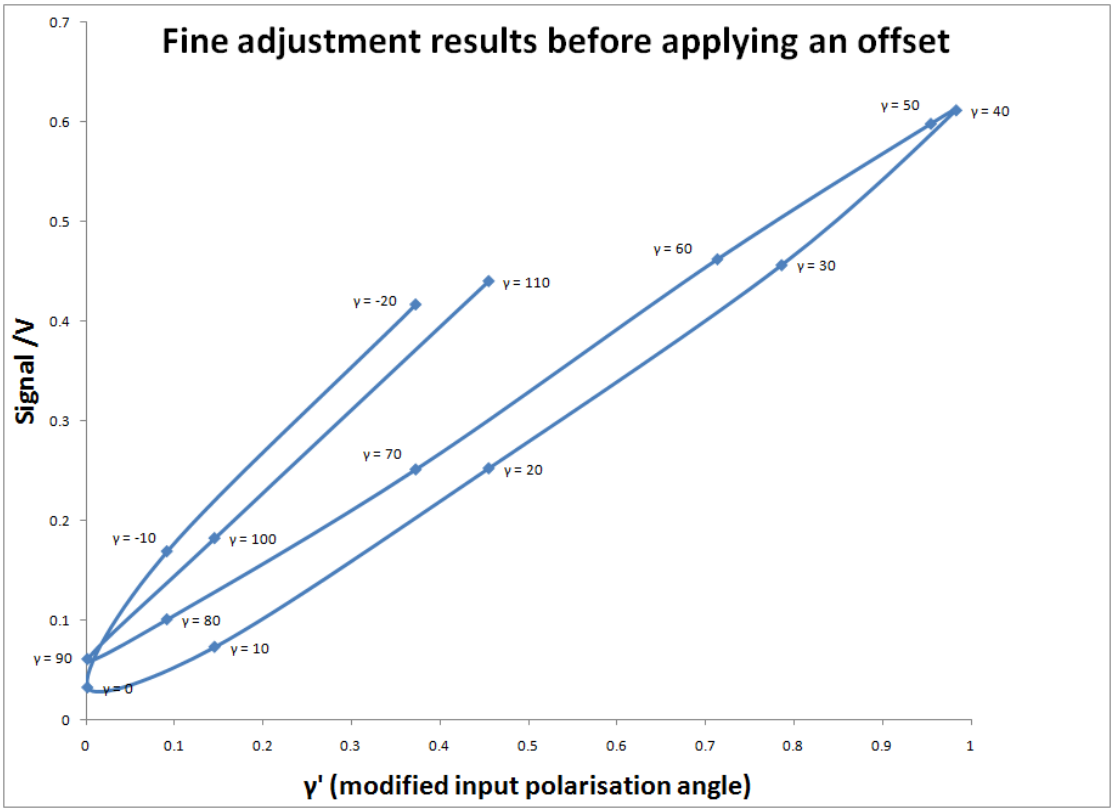


FIGURE 6.11: The fine adjustment experimental results before applying the offset, showing the system is not aligned

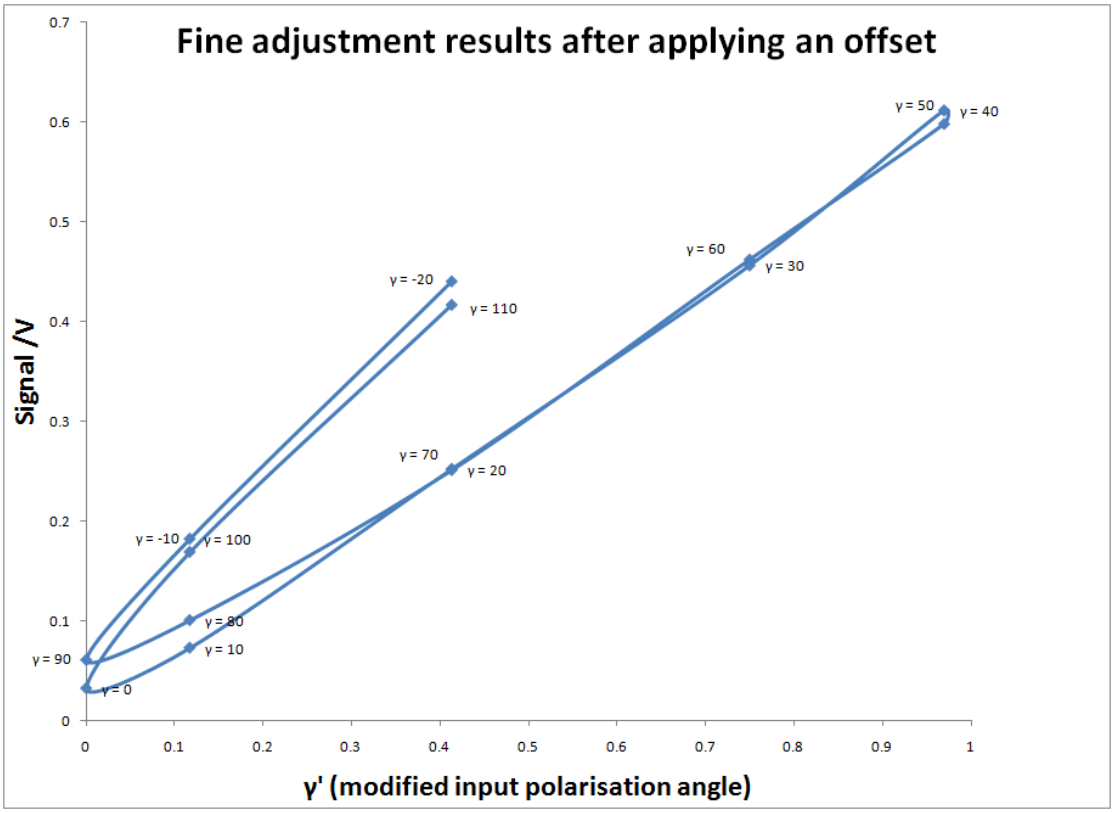


FIGURE 6.12: The fine adjustment experimental results after applying the offset, showing the system is now aligned within the experimental bounds

6.4 Subsidiary experiments

A number of subsidiary experiments were conducted to determine operational parameters and tolerances of the apparatus, providing additional metadata for the experiments and more trust in the data presented.

6.4.1 Laser power stability

As the intensity of the second harmonic signal is correlated to the power of the incident beam it is important to understand how the laser power varies with time. Should a constant drift occur the results would require normalising to a reference response and should a major fluctuation occur it may be necessary to re-run the experiment. The power stability was monitored at two times, the first as the laser was switched on and the second after a number of hours. The results are shown in Fig. 6.13 and 6.14.

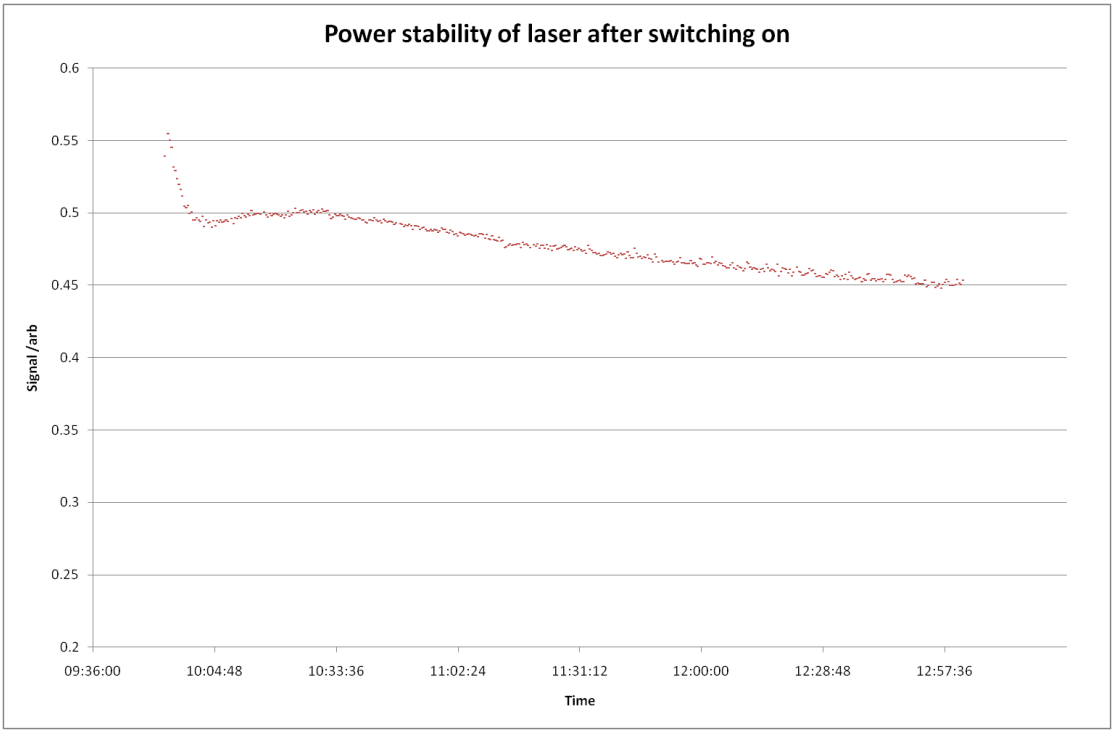


FIGURE 6.13: The power stability of the laser directly after switching on. Each point is the mean of a dataset

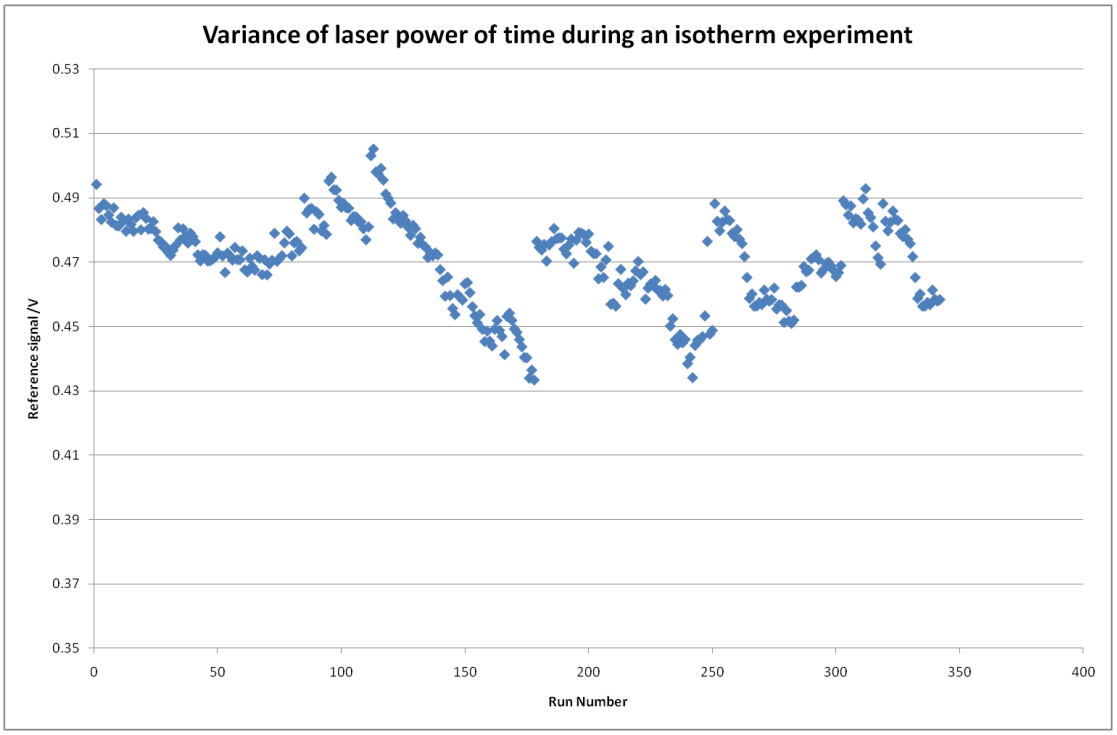


FIGURE 6.14: The power stability of the laser after it had been running for approximately 10 hours. Each point is the mean of a dataset

Fig. 6.13 shows that the laser can be considered 'ready' after approximately thirty minutes with a minor drift over the scale of an experiment. Fig. 6.14 shows reference data collected over a number of sequential experiments, covering approximately 5 hours. The variability is thought to be due to the opening and closing of the shutter, changing the temperature of the doubling crystal within the laser enclosure. To account for this variability and drift, the second harmonic signal will be normalised to laser power. As shown in Eq. 6.5 the second harmonic intensity is proportional in the square of the laser power, therefore the normalised signal will be calculated as:

$$I_{norm} = I_{SHG}/I_{ref}^2 \tag{6.50}$$

6.4.2 Monochromator scans

When the laser source is incident on a sample there are a number of other processes, such as two photon fluorescence, which may produce a signal at a wavelength other than that of the incident light. It is a possibility that the signal from these effects may overlap the wavelength of the second harmonic (266 nm), therefore a scan of monochromator wavelength at a given polarisation is completed. This provides information on effects taking place at the sample. An ideal response would show a sharp peak at 266 nm and no other effects. With the set up used in this work this was unattainable as the slit width selected on the monochromator resulted in it being less frequency specific,

broadening the peak. A monochromator scan was carried out for each compound used. An example from PNP is shown in Fig. 6.15. This shows a peak with a maximum at 266 nm, as expected, and no other effects close to this frequency. The trace show a non-zero baseline due to light 'leaking' into the detector, this will be removed from the background scans.

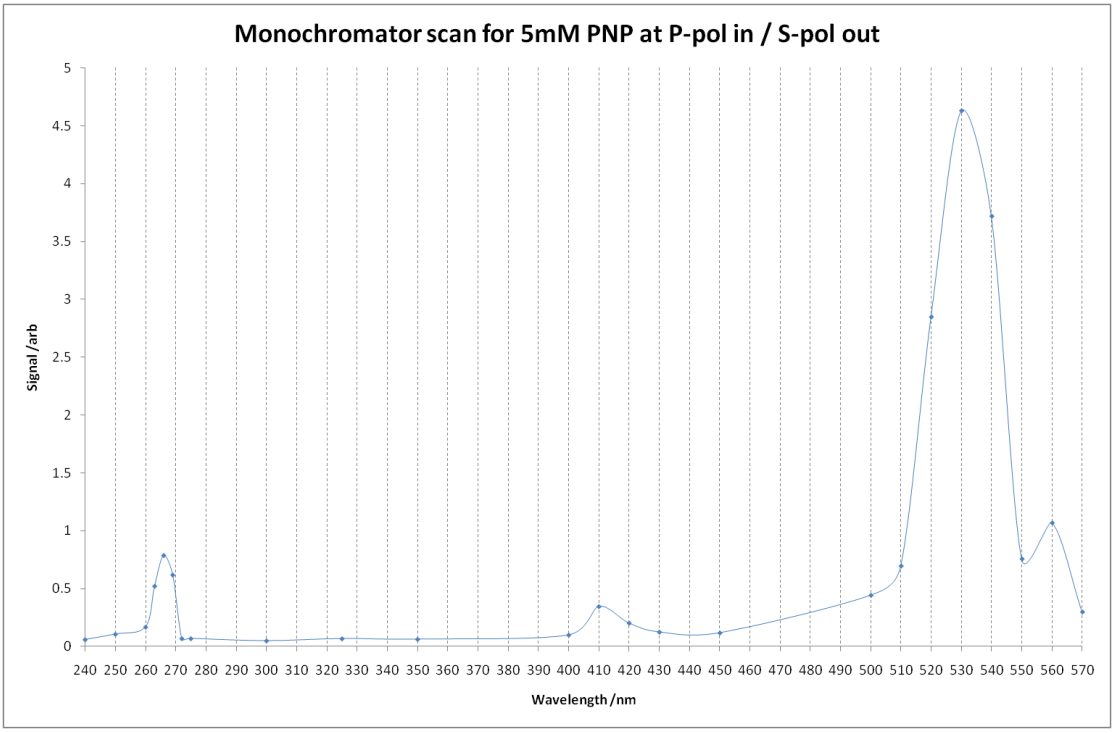


FIGURE 6.15: A monochromator scan for 5 mM PNP showing a peak at 266 nm, corresponding to the second harmonic, and no other effects close to this wavelength

6.4.3 Sample diluter accuracy

To automate the experimental process required to determine energy of adsorption to the surface a controllable SHG sample diluter was integrated into the laser rig. To determine the accuracy of the diluter an experiment was run as if an isotherm was being carried out. A sample of pure water was diluted by water through a number of 5ml sequential dilutions. Should the volume aspirated and/or dispensed from either syringe of the diluter not be equal the sample level would change and therefore the laser would be moved out of alignment. The results are shown in Fig. 6.16, indicating that the signal remained aligned for up to 25 dilutions. It is believed the signal decreases after this due to evaporation of the sample rather than an error from the diluter.

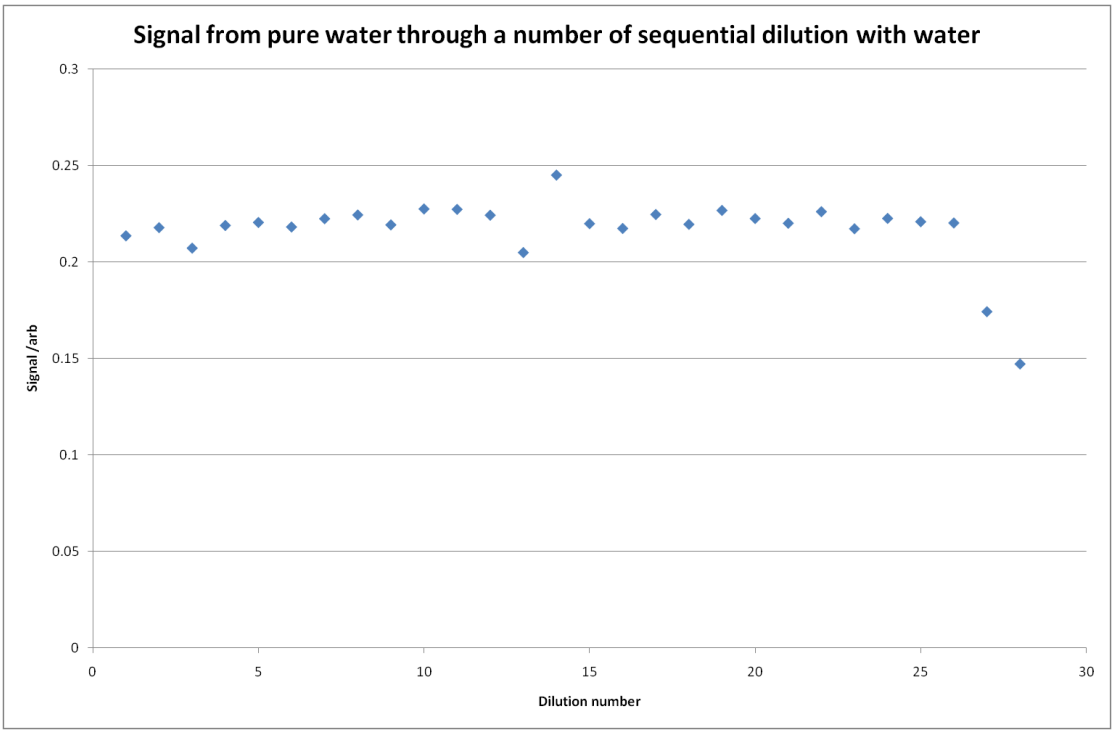


FIGURE 6.16: The results from 5 ml sequential dilutions of pure water with itself, showing no change in signal within 25 dilutions

6.4.4 Evaporation effects

As the second harmonic signal must pass through a number of irises and the monochromator slit in the output arm, any variations in alignment can have a dramatic effect on the observed intensity. As the beam travels at an angle of sixty degrees to the surface normal of the sample a shift in sample level will change the alignment of the system. As each pulse of the beam contains around 5 mJ of energy the sample may be heated and therefore increase the rate of evaporation. To obtain an understanding of this an experiment was run using a solution of pure water and running at a fixed polarisation. The results in Fig. 6.17 show that the evaporation was having a major effect of the results and therefore would need to be controlled.

To slow the evaporation a 'lid' was produced to fit over the sample, this was crafted out of a plastic powder funnel with slots cut for the input and output beam to travel through. Fig. 6.18 show this lid in-situ. To determine if this would slow the evaporation the original experiment was repeated. The results in Fig. 6.19 show that the rate of evaporation has been slowed to lower than the time-scale of an experiment.

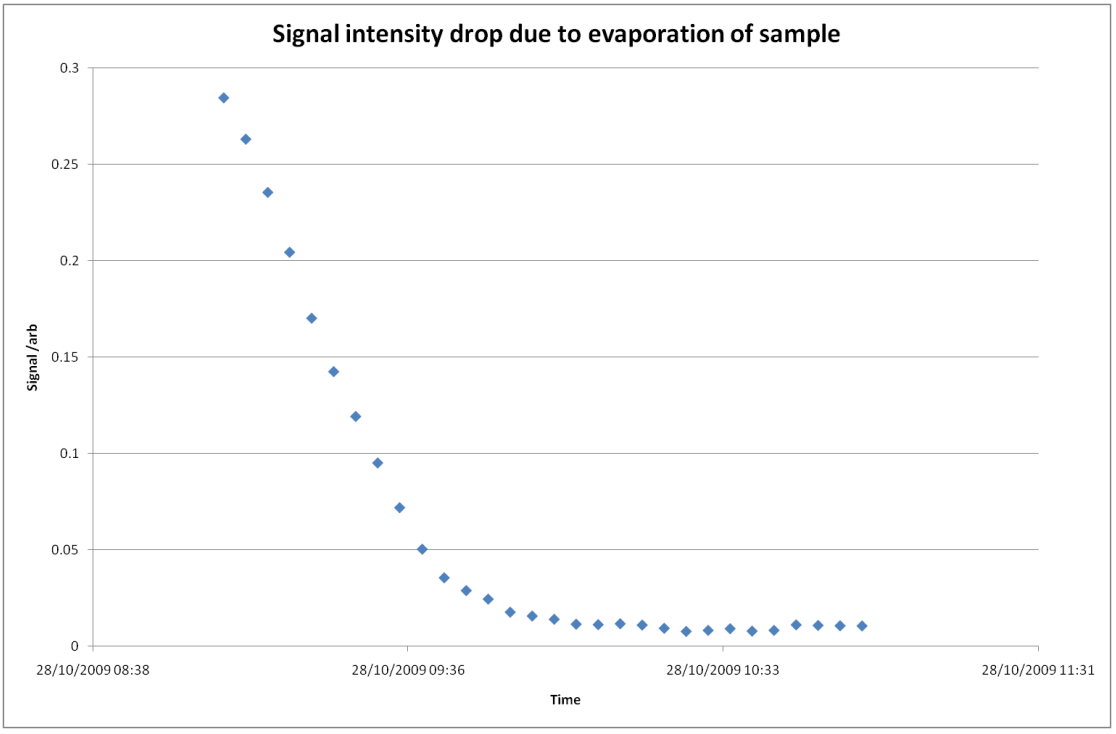


FIGURE 6.17: The drop in signal intensity due to evaporation of the sample, effecting the alignment

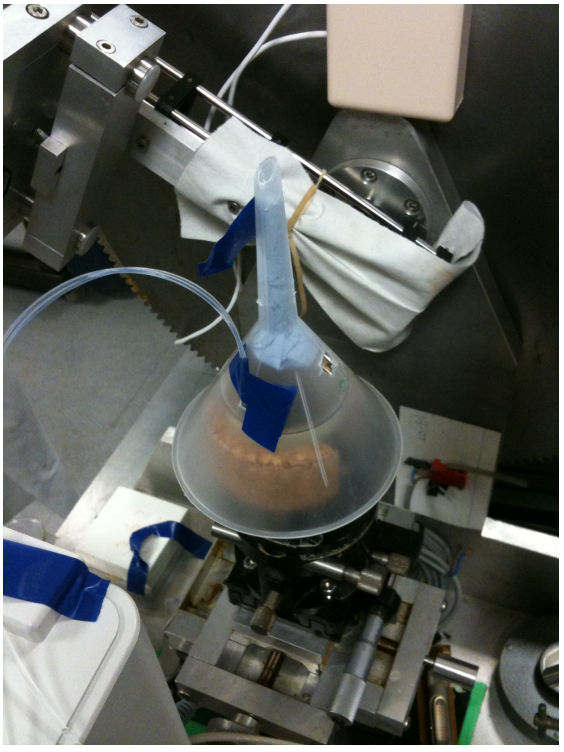


FIGURE 6.18: A photograph of the crafted lid in situ

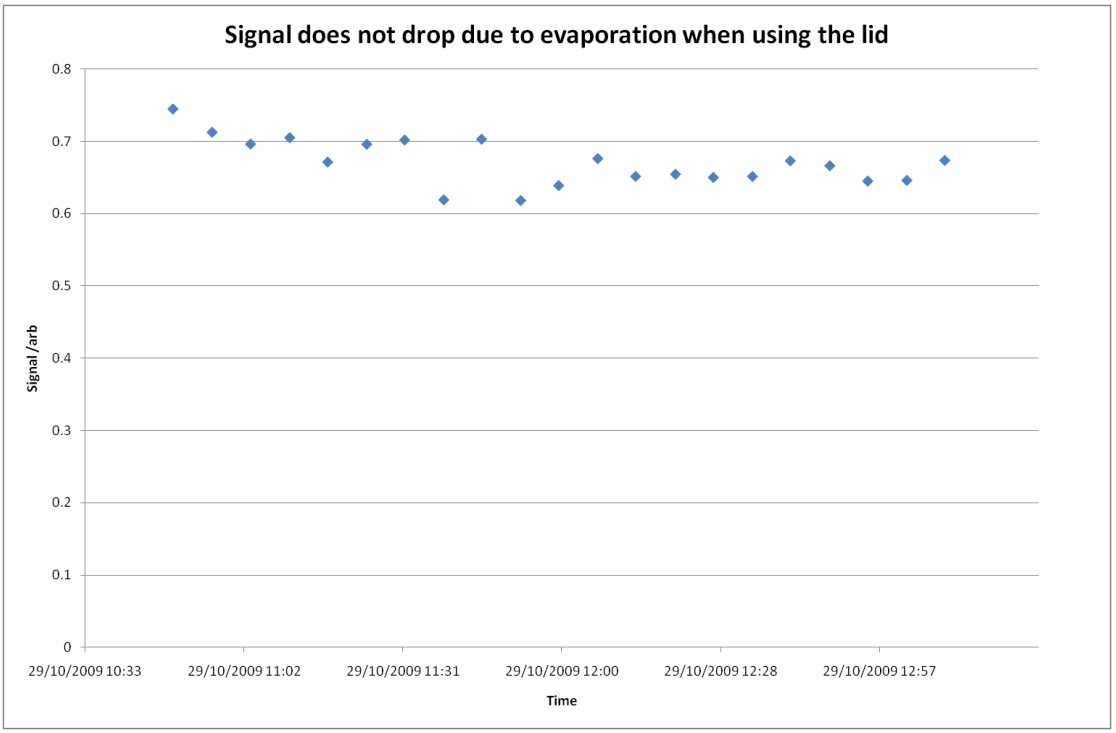


FIGURE 6.19: The results of running over a given time with the lid fitted, showing no drop in signal due to evaporation over the time scale of an experiment

6.4.5 Half wave plate tolerance

As the half wave plates were mounted in an automated rotator, moved via a stepper motor controller, the tolerances of the motors were determined. This allowed trust that the half wave plate would be moved accurately to the required position when requested. Some level of this checking was carried out when finding the true zero points of the half wave plates as these would be rotated by small increments until the optimum position was found. This was extended by carrying out an experiment with a fixed sample. Data was collected at two half wave plates positions, alternating between each over a number of runs. The results are shown in Fig. 6.20. The results show that over a number of scans the results remain unchanged within experimental error, therefore it will be assumed that the given half wave plate position is accurate and repeatable.

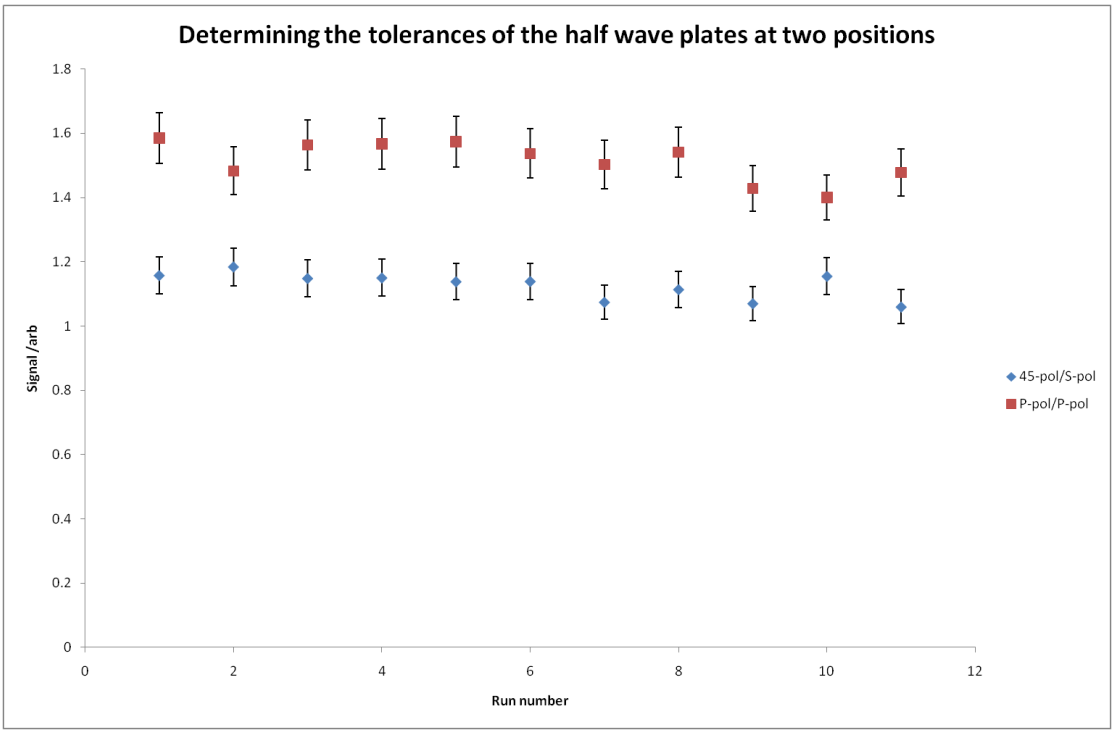


FIGURE 6.20: The results collected during the stepper motor tolerance experiments

6.4.6 Water bath calibration

A number of experiments will be run with the sample at a controlled temperature. This is achieved through use of a custom piece of glassware to contain the sample, which is linked to a water bath. The glassware contains a jacket of water around the sample and this water is pumped from and returned to the water bath.

An experiment was run to calibrate water bath setting to sample temperature. A waterproof temperature sensor was submerged within a sample of water and the water bath set to a given temperature. The laser was also active during these experiments as the beam may increase the temperature of the sample. The readings from the temperature sensors were recorded over an hour and the mean value determined. The results from this experiment are shown in table 6.2.

Water Bath Temp. /K	Sample Temp. /K	Variance of mean sample temp. /K
273.0	278.8	0.002
283.0	284.8	0.007
293.0	293.5	0.001
303.0	301.7	0.001
313.0	309.1	0.002
323.0	315.9	0.004

TABLE 6.2: Water bath temperature related to sample temperature

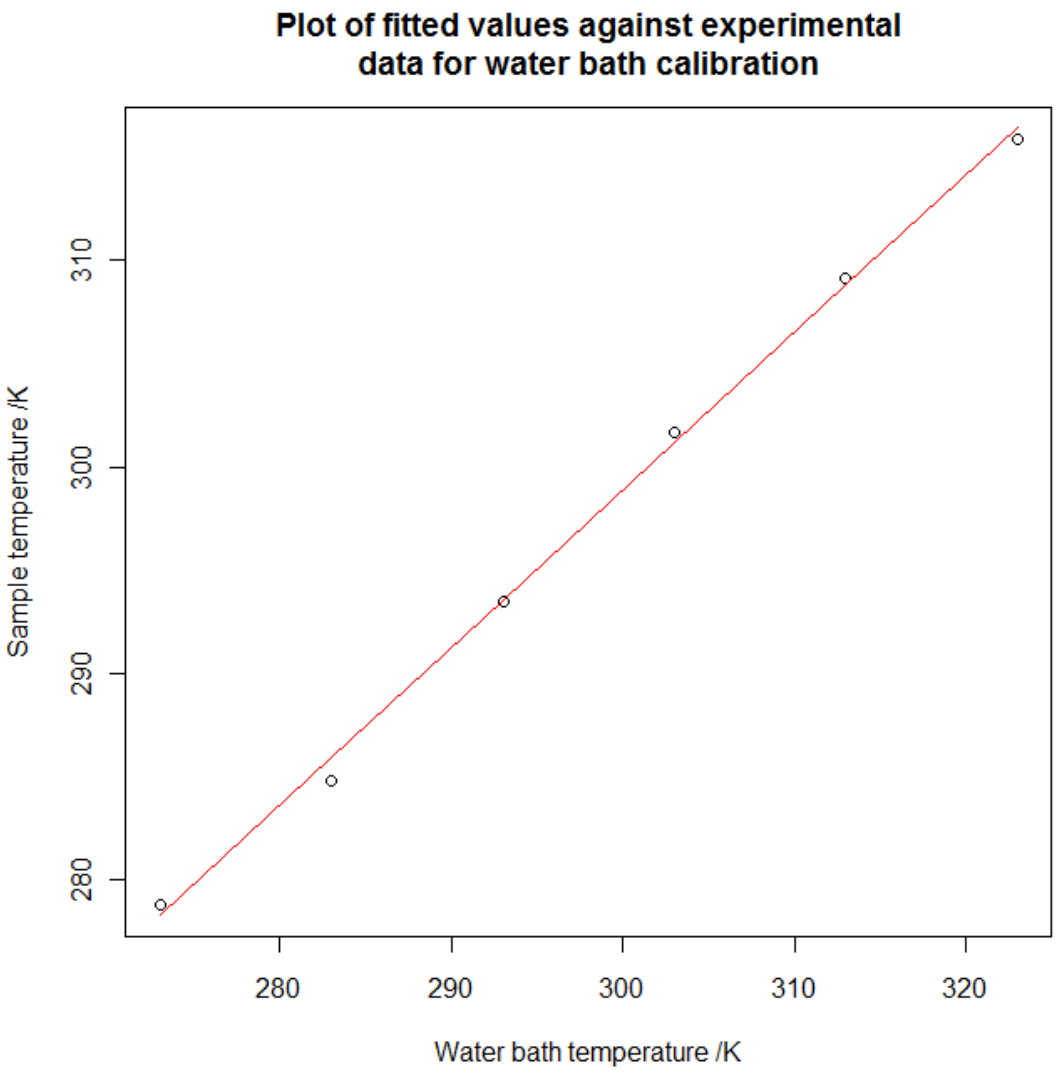


FIGURE 6.21: The fitted equation against experimental data for the water bath calibration

A linear regression was completed on this data to give the calibration equation shown in Eq. 6.51 and has been plotted with experimental values in Fig. 6.21.

$$Temp_{sample}/K = 0.762 * Temp_{waterbath}/K + 70.31 \tag{6.51}$$

6.5 Procedure

6.5.1 Sample preparation

To prevent the SHG signal from being modified by contamination all glassware was washed and prepared before each experiment. The samples were prepared to accurate

volumes in a volumetric flask and degassed by bubbling nitrogen through it for at least ten minutes. This was then transferred to a glass dish used in the experiments.

Before any piece of glassware was used it was irrigated with excess water and rinsed with acetone. This was then dried using a nitrogen flow. When switching to a new compound all glassware was washed with nitric acid and irrigated with excess water before going through the standard preparation procedure.

6.5.2 Experimental work-flow

Although there were a number of different types of SHG experiments carried out, they all followed a similar work-flow. On initiating an experiment the laser source would be polled to determine the laser status and if required start the laser. The system would then configure the multiple optics to the required parameters of the first scan and perform the data collection. The shutter would then be closed and another, shorter, data collect would be conducted to capture a background. The system would then determine if another scan was required and if so repeat the process, otherwise the laser would be shut down and the experiment complete. This process is discussed further in chapter 5. The specific work-flow of each experiment type is discussed below.

6.5.2.1 Polarisation scans

As shown in section 6.1.2, observing the intensity of the second harmonic at various polarisations can be used to determine the molecular orientation at the surface. This is performed experimentally by scanning the input polarisation between S-pol and P-pol for a number of output polarisations, typically S-pol, P-pol, 45^o and -45^o. These results are then fitted to a theoretical formula, which can be used to calculate the orientation. Fig. 6.22 shows the work-flow for this type of experiment.

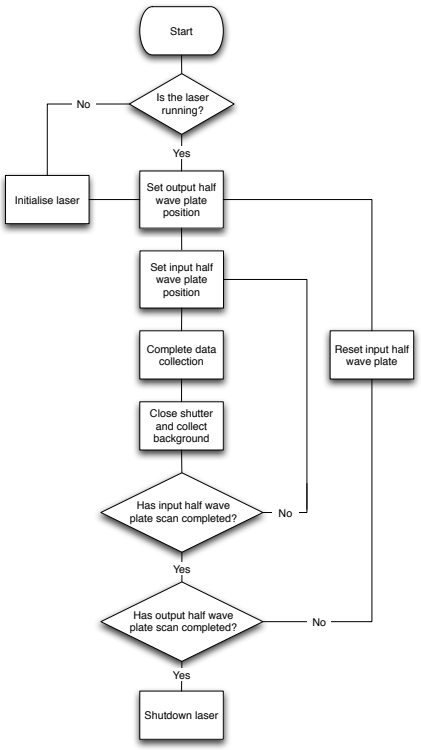


FIGURE 6.22: The work-flow for the polarisation scan experiments

6.5.2.2 Adsorption isotherms

As shown in section 6.1.3 by varying the concentration of the sample the adsorption equilibrium constant and adsorption free energy of the molecule to the surface can be determined. To perform this experiment a sample was prepared at the highest required concentration. The input and output polarisation was fixed to a point giving high signal, such as 45-pol input/S-pol output, and the sample scanned. This sample would then undergo a number of serial dilutions, with a scan at each concentration, until the signal has dropped to a value similar to that of the background. As there may be a contribution to the signal from water at lower concentrations a sample of pure water would be prepared. This would undergo the reverse of the previous experiment, having a number of serial concentrations from a bulk sample solution until enough data has been collected to provide a reasonable overlap of the two experiments. As shown above, evaporation has been an issue on longer experiments, therefore a typical experiment had ten concentration changes before the sample level was checked (and corrected). This work-flow is shown in Fig. 6.23.

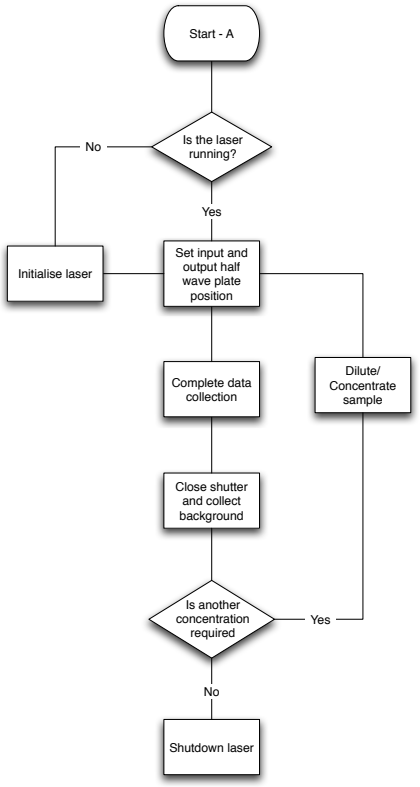


FIGURE 6.23: The work-flow for the adsorption isotherm experiments

6.5.2.3 Combined polarisation/adsorption isotherms

To increase the volume of data capture in one adsorption isotherm these were combined with the polarisation scans. These run as with the adsorption isotherms, capturing data before altering the concentration and repeating the process. In these experiments a number of polarisations were scanned at each concentration. Typically a full polarisation sweep was not conducted as the time scale of these experiment would mean that evaporation would greatly effect the observed results. The work-flow for these experiments is the same as shown in Fig. 6.23.

Chapter 7

Water

7.1 Introduction

As discussed previously, the SHG experiment is specific to the interfacial region. In this work the liquid/air interface will be studied. The liquids under investigation will be an organic compound dissolved into an aqueous solution, the results from these will be discussed later. As well as the desired compound, there will also be water molecules at the surface which will contribute to the second harmonic signal. It is therefore important to understand and characterise the signal obtained from pure water. When the compounds are at a reasonable concentration the signal from water will be small relative to that from the compound, but as concentration dependency experiments will be carried out the contribution from water will be important.

7.2 Experimental

The water used throughout the experiments was purified by a FiStreem Calypso Ultrapure water still and stored in a PTFE bottle. As surface contaminates will affect the results obtained from the SHG experiment, it is therefore important to know the purity of the sample. This was measure via surface tension, using a torsion balance.

7.2.1 Surface tension

The torsion balance experiments are carried out using a platinum ring suspended from the torsion balance. This ring is cleaned by heating with a flame until the metal glows white. The torsion balance is first calibrated to a zero point by aligning both the measurement and baseline needles to zero. The sample is then raised up until the ring has been submerged into the solution. As the ring is lifted from the surface the user

lowers the sample dish to maintain the baseline at the zero point. At the point when the ring pulls away from the surface, the surface tension value is recorded. Table 7.1 shows the surface tension from a number of water sources including that of the source used in the following experiments. These experiments were carried out at 293K.

Sample	Surface Tension /mNm^{-1}
Tap water	71.0
PURELAB Option S-R 7-15 (Surface Sciences)	72.5
FiStreem Calypso (Stulz group)	72.8
Literature[101]	72.75

TABLE 7.1: Surface tension results for a number of water sources collect at 293K, including the one used for further experiments

7.2.2 Second Harmonic Generation

The SHG signal at the water/air interface was recorded following the procedure described in Section 6.5.2.1. The results from this are shown in Fig. 7.1.

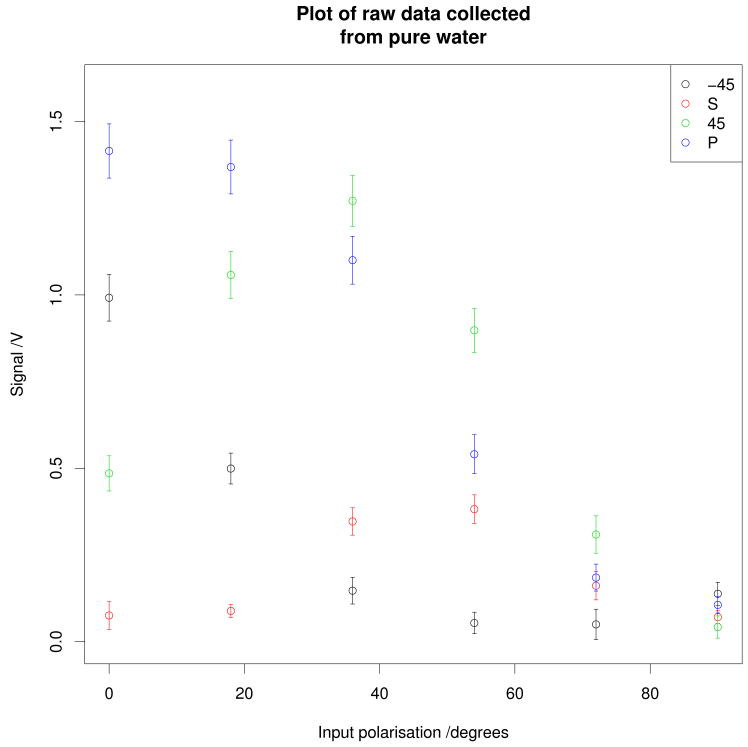


FIGURE 7.1: Normalised second harmonic signal as a function of input polarisation for water. Error bars are $\pm 3\sigma$

The second harmonic signal generated by water should contain no resonance enhancement component and as such the fitted values, A, B and C, should all be real (non-complex). Using the fitting equations and forcing A, B, and C to be real a measurement

of the surface purity can be obtained. If a poor fit is obtained than it can be assumed that either there is an alignment issue in the laser rig or the sample contains a contaminant. This fitting was completed on the above data and the results are shown in Fig. 7.2.

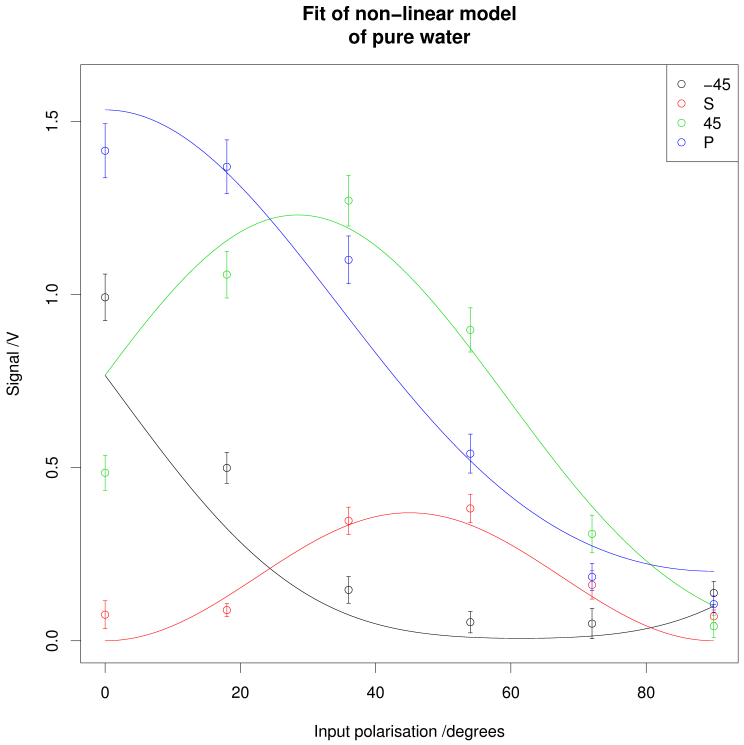


FIGURE 7.2: Non-linear fitted curves plotted with the experimental results collected for water. Error bars are $\pm 3\sigma$

This plot shows a poor fit of the water experimental data. Initially this was believed to be surface contaminates causing the deviation of the fit from the experimental values. After further investigation, it was found that by applying an offset to the output polarisation angle an improved fitting could be attained. In the case of the water sample a -12^o offset was applied, the plot from this fitting can be seen in Fig. 7.3. Although there is a good fit when A, B and C are fixed to real, the data was also fitted for allowing combinations of factors to be complex, the results from which are shown in table 7.2 and plots shown in Fig. 7.4

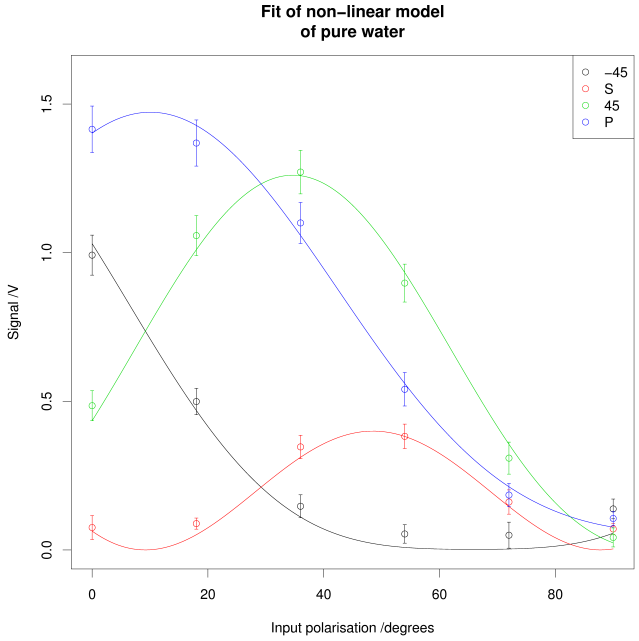


FIGURE 7.3: Non-linear fitted curves plotted with the experimental results collected for water after applying a -12^o offset to the output polarisation. Error bars are $\pm 3\sigma$

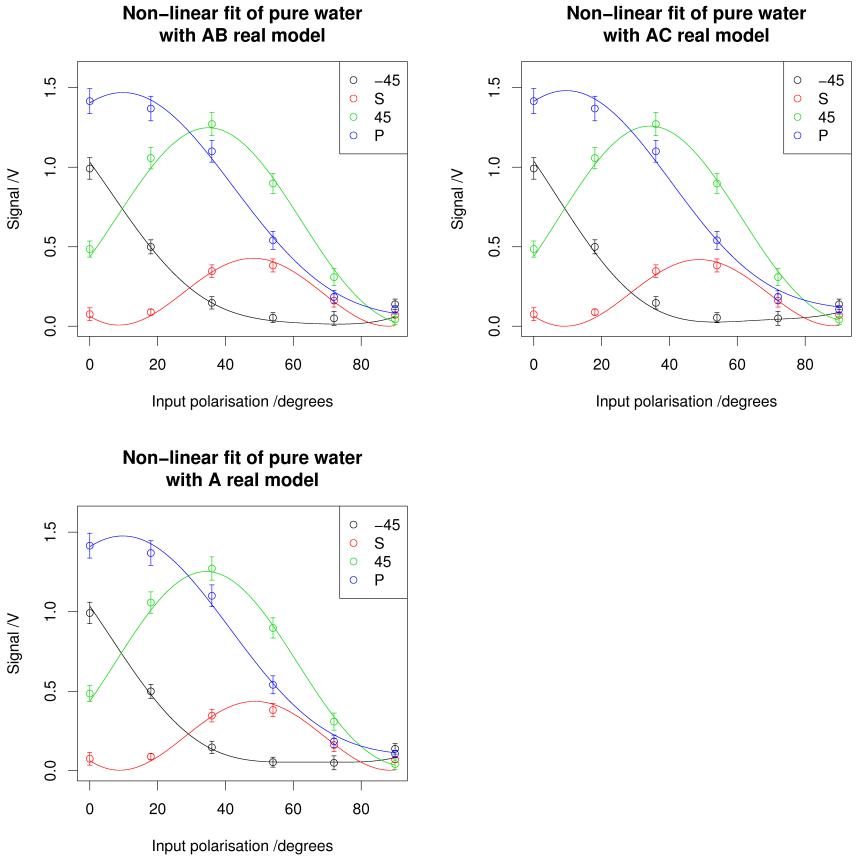


FIGURE 7.4: Non-linear fitting lines plotted with the experimental results collected for water after applying a -12^o offset to the output polarisation. Error bars are $\pm 3\sigma$

Model	A	B	C	RSS	ANOVA	
					F test	Pr(>F)
ABC real	1.211±0.011	0.283(±0.025)	-0.799(±0.015)	0.0383		
AB real	1.211±0.01	0.292(±0.024)	-0.762(±-0.039) + 0.278(±0.074)i	0.0334	2.977	0.100
AC real	1.216±0.01	0.224(±0.031) + 0.277(±0.063)i	-0.807(±0.014)	0.0305	5.144	0.035
A real	1.213±0.01	0.262(±0.035) + 0.223(±0.079)i	-0.791(±0.03) + 0.187(±0.092)i	0.0265	4.892	0.040

TABLE 7.2: Non-linear fitting values determined from the water data after applying a -12^o offset to the output polarisation (RSS - Residual sum of squares)

Review of this data shows that the fitting when only A is real gives the lowest residual sum of squares. Although this may suggest this is the optimal fit, the reduced error may also be a factor of the extra freedom in the fitting model. It can also be seen in the fits where the factors are complex, the imaginary parts are close to zero. The results from the ANOVA test show that the significance of adding the extra degrees of freedom (by making B and/or C complex) are low, and as such have little effect on the fitting.

These results were converted to the susceptibility terms, χ, using Eqs. 6.16 6.17 and 6.18. The ratio between these terms are shown in table 7.3 and compared to those determined in previous work.

Data	χ_{zzz}/χ_{zxx}	χ_{zzz}/χ_{xxz}	χ_{zxx}/χ_{xxz}
This work	7.48 ± 0.812	3.60 ± 0.103	0.480 ± 0.053
Haslam *et al.*[102]	7.72	3.28	0.42
Fordyce *et al.*[103]	7.11	3.37	0.47
Timson *et al.*[104]	6.49	3.33	0.51
Tamburello-Luca *et al.*[105]	6.20	2.55	0.41
Goh *et al.*[106]	2.60	1.20	0.46
Sokhan *et al.*[107]	7.17	5.91	0.82
Antoine *et al.*[108]	3.72	2.08	0.56

TABLE 7.3: The ratio between χ terms for the A, B and C real fit of the water data, including ratios determined in previous work

To calculate these susceptibility terms an estimate of the refractive indices of the interface at fundamental and second harmonic wavelengths is required. It has been shown by Fordyce *et al.*[103] how the assumptions of the refractive indices greatly effect the determined ratios. In this work the values used were those reported by Fordyce *et al.* to allow comparison. These assumptions are that the refractive indices at both fundamental and second harmonic wavelength is the average between the two bulk media. The calculated susceptibility ratios show a good fit and suggest the SHG system is correctly aligned and producing accurate results. The results reported in the other work either use refractive indices of pure water for that of the interface or are not reported and may explain the deviation between results.

These results can be further analysed to determined the molecular orientation parameters, D, and the ratio between the dominant hyperpolarisability terms, β_{zxx} and β_{zzz}. To undertake this fitting process a better estimate of the refractive indices at fundamental and second harmonic wavelength is required for the interface, n_3. In previous studies it has been assumed this is equal to that of bulk water. It has been shown by Greef *et at.*[109] that the water surface is more dense than in the bulk and therefore has a higher refractive index. This was determined using spectroscopic ellipsometry. From this work the refractive indices for water were determined as 1.339 and 1.374 for fundamental and

second harmonic wavelength respectively. Using these values the water data was fitted to Eq. 6.28 and 6.29. The orientation parameter, D, was determined as 0.805 $\pm$ 0.0034 which correspond to an average angle of 26.19^o from surface normal. The ratio between dominant hyperpolarisability terms was determined as -0.112 $\pm$ 0.0094.

7.3 Discussion

It has been shown that with the implemented optical and electrical alignment discussed in chapter 6, a SHG signal can be obtained at the air/water interface. These results are reliable and repeatable with a high degree of accuracy. It has been shown that the water used in this study has a high purity through analysis of both surface tension and SHG. It has also been shown that through a combination of a Faraday Modulator, discussed in section 6.3.1, a high degree of optical alignment can be obtained for the optics within the system. The tweaking experiment, section 6.3.2, also allows accurate alignment of the input (523 nm) half-wave plate. The analysis of these results suggests that neither of these methods provide detailed information of the output (266 nm) half-wave plate. When completing further analysis of the SHG data this will be considered.

Chapter 8

Para-nitrophenol

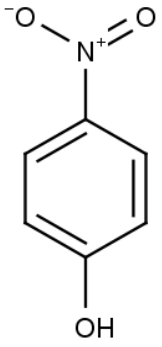


FIGURE 8.1: The molecular structure of PNP

PNP was chosen as a target molecule as it has a large hyperpolarisability, due to the electron donating and accepting groups either end of the π system. As the second harmonic signal is a measure of the hyperpolarisability, it should therefore give a high signal. PNP at the air/liquid interface has been studied via SHG in the past[110, 25] and therefore will allow a comparison of the control and review system developed as part of this work with previous results.

As well as providing a good test compound, PNP is also of interest in the study of pollution[111, 112, 113]. These types of compounds are used in many industrial processes and as such can be released into the environment. As these compounds are toxic to plant life and aquatic organisms it is important to understand the processes the compound can undergo. Through study of the PNP surface, a better understanding of interface and transfer processes can be gained.

8.1 Previous studies

There have been a number of previous studies on PNP at interfaces via surface specific techniques. These studies are summarised in table 8.1, 8.2 and 8.3.

Researcher	Interface	Determined parameters
Paluch *et al.*[114]	Air/Water	Δ_{abs}G at 20^{o}C = -10.51 kJ mol^{-1}
	Surface tension (Maximum bubble)	Δ_{abs}G at 30^{o}C = -10.55 kJ mol^{-1}
	fitted to Volmer isotherm	Δ_{abs}G at 40^{o}C = -10.59 kJ mol^{-1}
Tamburello-Luca *et al.*[115]	Air/water	Orientation parameter, D = 0.45
		Δ_{abs}G = -16.4 kJ mol^{-1}
	Hexane/water	Orientation parameter, D = 0.50
		Δ_{abs}G = -17.2 kJ mol^{-1}
Higgins *et al.*[116]	Silica/Air	Δ_{abs}G = -17 kJ mol^{-1}
		$\frac{\beta_{xzz}}{\beta_{zzz}}$ = -0.31
		Orientation parameter, D = 0.33
	Water/Air	Δ_{abs}G = -19 kJ mol^{-1}
		Orientation parameter, D = 0.26
	Chloroform/Silica	Orientation parameter, D = 0.27
Bell *et at.*[110]	Air/water (assumed layer above water)	$\frac{\beta_{xzz}}{\beta_{zzz}}$ = -0.25
		Orientation parameter, D = 0.40
	Air/water (assumed layer within water)	$\frac{\beta_{xzz}}{\beta_{zzz}}$ = -0.15
		Orientation parameter, D = 0.63
	Heptane/water (assumed layer above water)	$\frac{\beta_{xzz}}{\beta_{zzz}}$ = -0.11
		Orientation parameter, D = 0.61
	Heptane/water (assumed layer within water)	$\frac{\beta_{xzz}}{\beta_{zzz}}$ = -0.13
		Orientation parameter, D = 0.57

TABLE 8.1: A summary of the previous work complete on PNP at the interface. Note that the Tamburello-Luca Δ_{abs}G results are calculated from refitting the Paluch surface tension data to a Frumkin isotherm

Researcher	Interface	Determined parameters
Bhattacharyya *et al.*[117]	Air/water	Shows the pH dependence of SHG Suggests anionic form is not present at the surface
Haslam *et al.*[118]	Air/Water	Δ_{abs}G = -24.3 kJ mol^{-1} (Langmuir isotherm)
	Hexane/water (Surface tension measurements)	Δ_{abs}G = -17.2 kJ mol^{-1} (Frumkin isotherm)
	Heptane/water	Δ_{abs}G = -25.7 kJ mol^{-1} (Langmuir isotherm)
	Ethanol/silica	Δ_{abs}G = -27.3 kJ mol^{-1} (Langmuir isotherm)
	Water/glass	Δ_{abs}G = -14.1 kJ mol^{-1} (Langmuir isotherm)
	Dodecane/water	Δ_{abs}G = -31.4 kJ mol^{-1} (Langmuir isotherm) Have also studied the concentration dependence of Δ_{abs}G with the addition of TBP
Castro *et al.*[119]	Air/water	Δ_{abs}G = -21.5 kJ mol^{-1}
	Glass/water	Δ_{abs}G = -14.2 kJ mol^{-1}
Das *et al.*[120]	Air/water	Orientation parameter, $D = 0.60$ Addition of 5 M $LiClO_4$ - orientation parameter, $D = 0.60$ Addition of 5 M GdmCl - orientation parameter, $D = 0.65$ Addition of 5 M LiCl - orientation parameter, $D = 0.64$
		Δ_{abs}G = -21.6 kJ mol^{-1} Addition of 3 M $LiClO_4$ - Δ_{abs}G = -20.9 kJ mol^{-1} Addition of 3 M GdmCl - Δ_{abs}G = -21.0 kJ mol^{-1} Addition of 3 M LiCl - Δ_{abs}G = -23.5 kJ mol^{-1}

TABLE 8.2: Continuation of the summary of the previous work complete on PNP at the interface. TBP - tri-n-butyl phosphate

Researcher	Interface	Determined parameters
Steel *at al.*[121]	Water/cyclohexane	Orientation parameter, $D = 0.33$
Sarkar *et al.*[122]	Water/air	Orientation parameter, $D = 0.60$
		Addition of 5 M urea - orientation parameter, $D = 0.60$
		Addition of 0.2 mM CTAB - orientation parameter, $D = 0.64$
		Addition of 0.2 mM CTAB + 7 M urea - orientation parameter, $D = 0.65$
		Δ_{abs}G = -21.6 kJ mol^{-1}
		Addition of 1 M urea - Δ_{abs}G = -19.5 kJ mol^{-1}
		Addition of 3 M urea - Δ_{abs}G = -18.7 kJ mol^{-1}
		Addition of 0.1 mM CTAB - Δ_{abs}G = -21.6 kJ mol^{-1}
		Addition of 0.1 mM CTAB + 3 M urea - Δ_{abs}G = -19.1 kJ mol^{-1}
Timson *et al.*[104]	Air/water	orientation parameter, $D = 0.61$
	Dodecane/water	orientation parameter, $D = 0.74$

TABLE 8.3: Continuation of the summary of the previous work complete on PNP at the interface. CTAB - cetyltrimethylammonium bromide

8.2 Experimental

A HPLC grade (≥99.5% purity) sample of PNP was purified before use in the following experiments. A sample was dissolved in minimal hot water (70^{o}C). The solution was vacuum filtered hot through filter paper. The collected filtrate was allowed to cool to room temperature and filtered under vacuum through filter paper. The white crystals collected on the filter paper was washed with cold water (approx. 0^{o}C) and dried under vacuum. A sample was dissolved in deuterated acetonitrile and submitted to NMR. The NMR spectra is shown in Fig. 8.2 and shows no impurities of a high enough concentration to be detected by NMR.

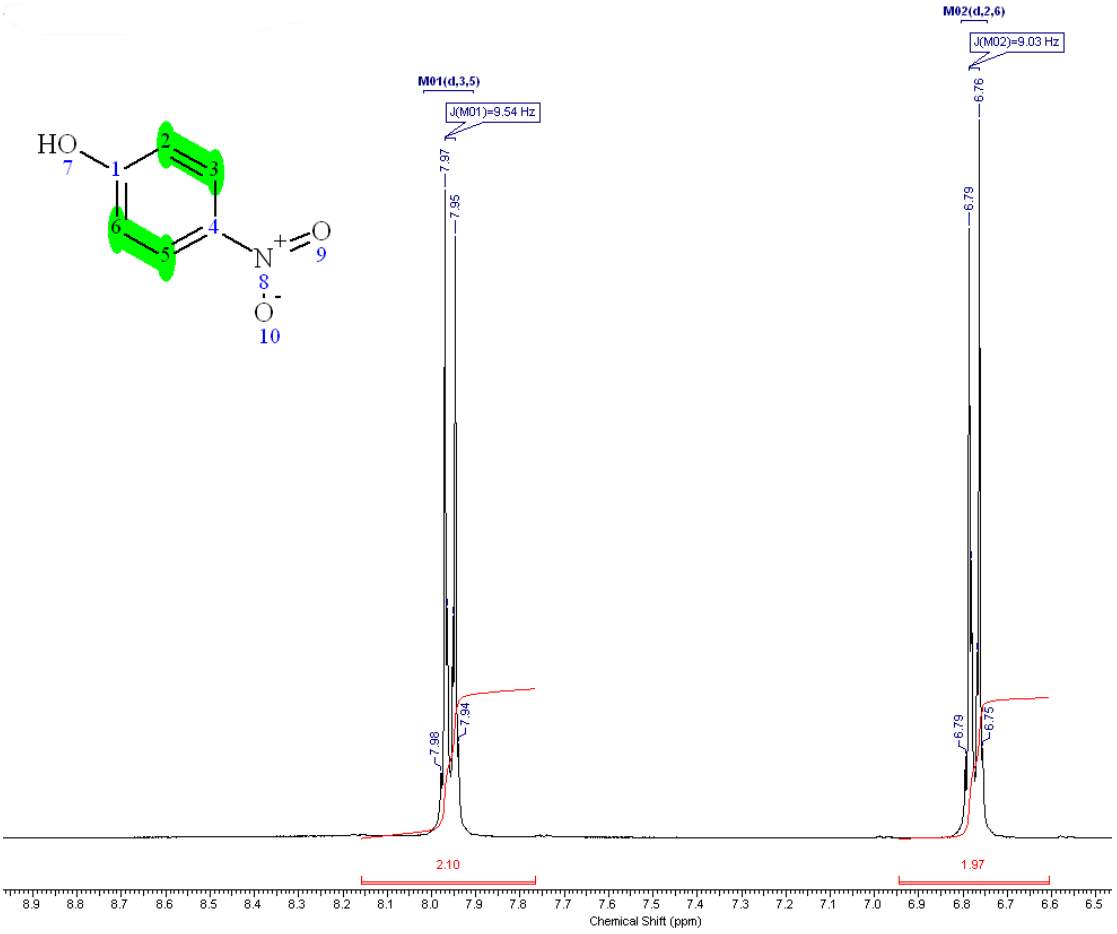


FIGURE 8.2: NMR spectra of purified PNP in deuterated acetonitrile

8.2.1 Temperature dependence

A study of the temperature dependence of the surface orientation was completed for PNP. The study investigated orientation changes which occur between 0^{o}C and 50^oC. These experiments were conducted using the water bath described in section 6.4.6 and followed the work-flow described in section 6.5.2.1. An initial set of experiments were

completed with water bath temperatures between 0^oC and 50^oC in 10^oC steps. The results from this indicated further resolution was required between 20^oC and 30^oC and therefore an additional set of experiments were carried out between these temperatures in 2^oC steps. These experiments were completed with a 75 mM solution prepared following the procedure described in section 6.5.1.

8.2.2 Concentration dependence

Initial investigation into the results from the temperature dependence experiments suggested the orientation of the molecule at the interface was also affected by concentration. To better understand this, a set of experiments were carried out at multiple concentrations. The experiments were completed following the work-flow described in section 6.5.2.1 and the water bath set to a constant of 20^oC . The concentrations investigated were 75, 50, 30 and 10 mM and prepared following the procedure described in section 6.5.1.

8.2.3 Adsorption isotherms

To determine the temperature dependence of the surface equilibrium constant, k, of PNP a number of adsorption isotherms were completed at water bath temperatures between 0^oC and 50^oC in 10^oC steps. The initial concentration of the sample was 75 mM and diluted through a number of serial dilutions to approximately 0.5 mM. A sample of pure water was then prepared and concentrated through a number of serial additions of a 5 mM stock solution until a concentration of approximately 1 mM. The two experiments were conducted this way to provide sufficient overlap of the results. The experiments were completed following the work-flow described in section 6.5.2.3, the polarisation combinations investigated were 45^o input/S output and S input/P output (during the 0^oC and 20^oC experiments 54 input/45 output was also included). These polarisation combinations were chosen as they correspond to a large signal observed with the PNP polarisation scan experiments. The 45 input / S output should show a significant contribution from water at lower concentrations whereas P input / P output should not as very little water signal is seen at this polarisation, as seen in Fig. 7.1.

8.3 Results

8.3.1 Temperature dependence

Due to evaporation issues at the higher temperatures a complete scan had to be conducted over a number of experiments. This allowed for re-alignment between the experiments. Table 8.4 describes which data was used to generate the initial dataset, 0^oC -

50^oC in 10^oC steps, used in the fitting software (referred to as 0-50 dataset). Table 8.5 describes the second dataset looking at high resolution between 20^oC and 30^oC (referred to at 20-30 dataset).

Water bath temp. /^oC	Experiment number	Experimental data
0	321	S output P output 45 output -45 output
10	355	S output P output 45 output -45 output
20	317	S output P output, 36^o to 90^o input 45 output, 0^o to 78^o input
	318	P output, 0^o to 30^o input -45 output
30	348	S output P output 45 output -45 output
40	319	S output, 18^o to 90^o input P output 45 output, 48^o to 90^o input -45 output
	320	S output, 0^o to 12^o input 45 output, 0^o to 42^o
50	373	S output
	376	P output
	375	45 output
	377	-45 output

TABLE 8.4: The combination of experiments used to generate the 0-50^oC temperature dependence of surface orientation dataset for 75 mM PNP. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

Water bath temp. /^{o}C	Experiment number	Experimental data
20	597	S output
		P output
		45 output
		-45 output
22	583	S output
		P output
		45 output
		-45 output
24	577	S output
		P output
		45 output
		-45 output
26	590	S output
		P output
		45 output
		-45 output
28	593	S output
		P output
		45 output
		-45 output
30	594	S output
		P output
		45 output
		-45 output

TABLE 8.5: The combination of experiments used to generate the second, more detailed, dataset used to determine the temperature dependence of surface orientation for 75 mM PNP. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

These datasets were fitted via non-linear regression to Eq. 6.21 using an R[89] script. The model was set up to consider A real with B and C complex. The fitting results for the 0-50 dataset can be seen in table 8.6 and a summary plot of the fitted curves can be seen in Fig. 8.3. As with the water data, an offset was applied to the output polarisation. As the system has been re-aligned between the water and PNP experiments the optimal offset for these datasets was $+3^o$. The fitting results from the 20-30 dataset can be seen in table 8.7 and a summary plot of the fitted curves can be seen in Fig. 8.4.

Water bath temp. /^{o}C	A	B	C	Residual standard error
0	1.142±0.033	-1.530(±0.110) - 1.996(±0.080)i	1.673(±0.073) + 1.489(±0.077)i	0.179
10	0.912±0.043	-1.387(±0.157) - 1.669(±0.123)i	1.309(±0.111) + 1.512(±0.093)i	0.183
20	1.218±0.050	-1.859(±0.172) - 2.392(±0.127)i	1.963(±0.124) + 2.145(±0.109)i	0.284
30	0.962±0.057	-1.523(±0.187) - 2.315(±0.118)i	1.901(±0.133) + 1.732(±0.139)i	0.210
40	0.954±0.089	-1.303(±0.314) + 2.858(±0.135)i	1.297(±0.216) - 2.609(±0.105)i	0.395
50	0.693±0.078	-2.343(±0.415) - 1.379(±0.689)i	2.453(±0.177) + 0.526(±0.782)i	0.332

TABLE 8.6: Non-linear fitting values determined from the 0-50^{o}C 75 mM PNP dataset after applying a +3^o offset to the output polarisation

Water bath temp. /^{o}C	A	B	C	Residual standard error
20	0.362±0.024	-0.416(±0.086) - 0.665(±0.052)i	0.549(±0.056) + 0.485(±0.059)i	0.027
22	0.327±0.029	-0.262(±0.105) - 0.665(±0.040)i	0.502(±0.069) + 0.502(±0.064)i	0.031
24	0.411±0.022	-0.482(±0.077) + 0.689(±0.052)i	0.694(±0.053) - 0.559(±0.062)i	0.028
26	0.385±0.048	-0.512(±0.197) + 0.860(±0.111)i	0.473(±0.133) - 0.784(±0.079)i	0.056
28	0.381±0.044	-0.429(±0.175) + 0.913(±0.078)i	0.428(±0.114) - 0.820(±0.058)i	0.050
30	0.380±0.033	-0.387(±0.121) + 0.893(±0.050)i	0.530(±0.080) - 0.736(±0.054)i	0.039

TABLE 8.7: Non-linear fitting values determined from the 20-30^{o}C 75 mM PNP dataset after applying a +3^o offset to the output polarisation

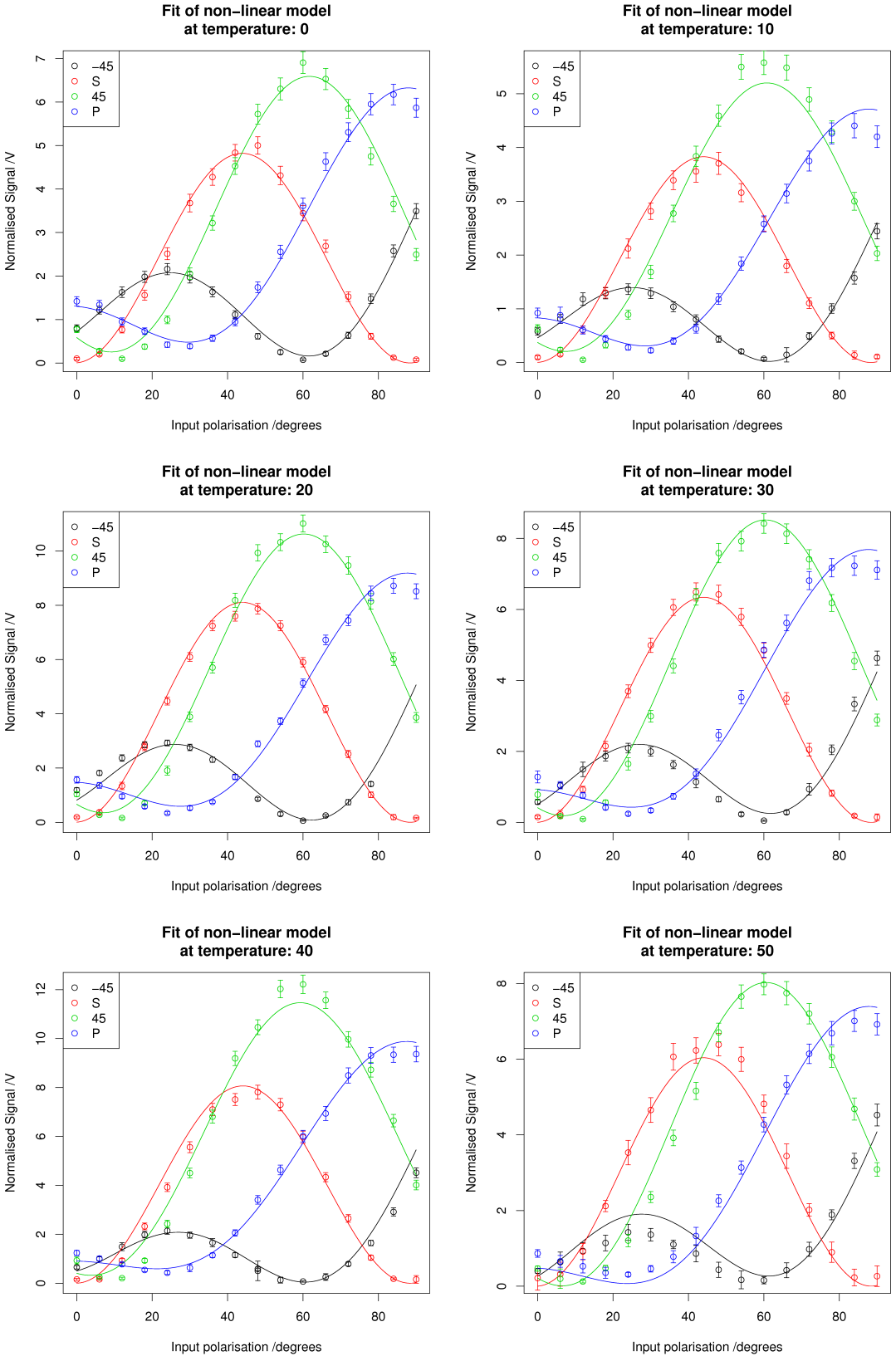


FIGURE 8.3: The summary plot of fitted curves against experimental data for each temperature from 0-50^{o}C 75 mM PNP dataset after applying the +3^o offset to the output polarisation. Water bath temperatures are given. Error bars are $\pm 3\sigma$

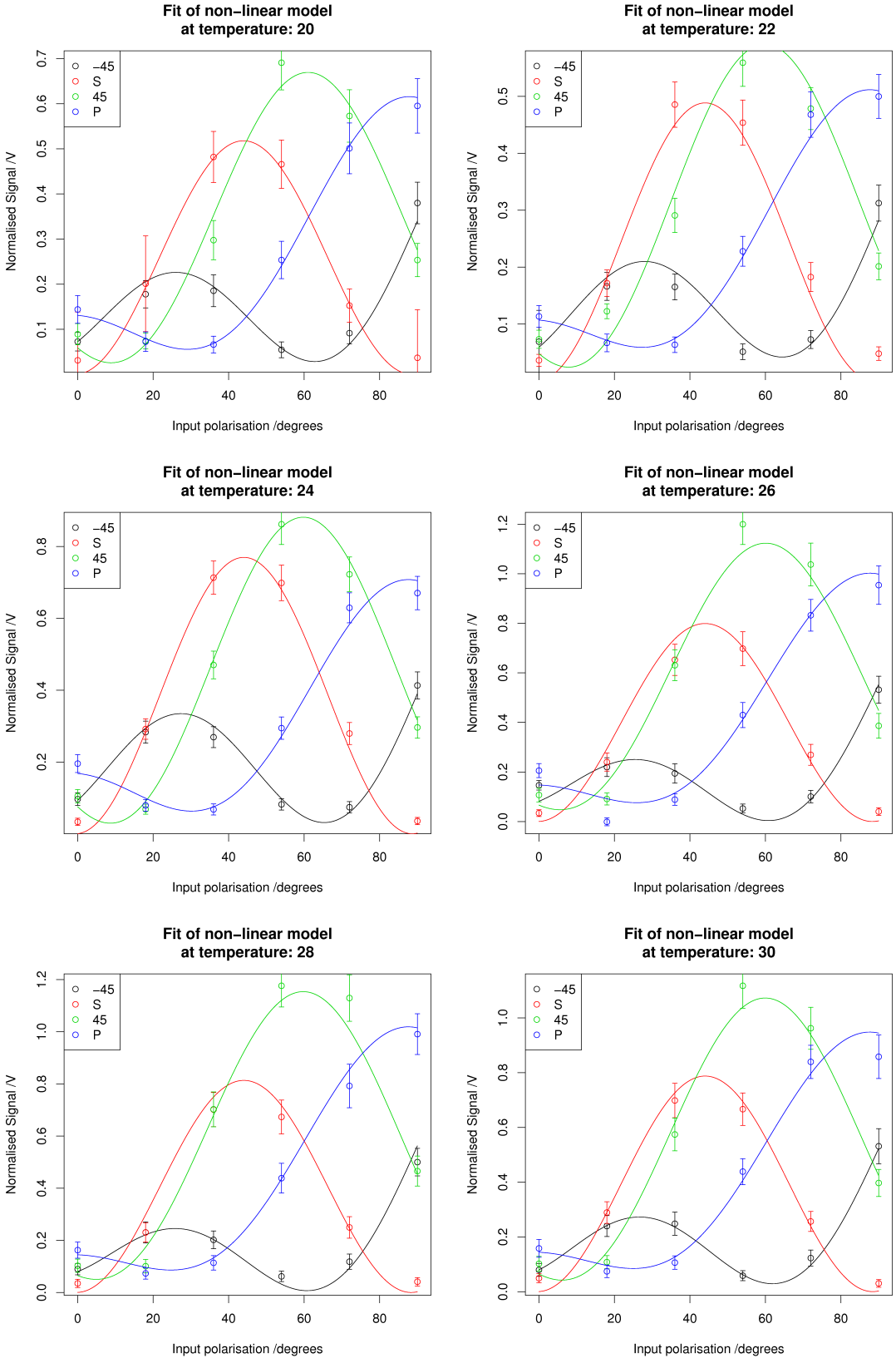


FIGURE 8.4: The summary plot of fitted curves against experimental data for each temperature from 20-30^{o}C 75 mM PNP dataset after applying the $+3^o$ offset to the output polarisation. Water bath temperatures are given. Error bars are $\pm 3\sigma$

These fits were verified by modifying the fitting equation, Eq. 6.21, to replace the $\cos(\phi)$ and $\sin(\phi)$ terms with t functions where:

$$t = \tan(\phi/2) \tag{8.1}$$

$$\sin(\phi) = \frac{2t}{1+t^2} \tag{8.2}$$

$$\cos(\phi) = \frac{1-t^2}{1+t^2} \tag{8.3}$$

This conversion applies a constraint to ϕ to keep it within mathematical limits, while keeping the overall model unconstrained. The fitting equation when considering B and C complex then becomes:

$$\begin{aligned} I_{SHG} = [(rC\frac{1-t_c^2}{1+t_c^2}\sin(2\gamma)sin(\Gamma) + (A\cos^2(\gamma) + rB\frac{1-t_b^2}{1+t_b^2}\sin^2(\gamma))cos(\Gamma)]^2 \\ +[rC\frac{2t_c}{1+t_c^2}\sin(2\gamma)sin(\Gamma) + (rB\frac{2t_b}{1+t_b^2}\sin^2(\gamma))cos(\Gamma)]^2 \end{aligned} \tag{8.4}$$

The R script was modified to include the new equation and the two 75 mM PNP dataset were re-fitted. A plot of the 20^{o}C fits from both 75 mM PNP datasets can be seen in Fig. 8.5 and table 8.8 shows the fitted values using both equations. These results show no variation between the two models, validating the fitting procedure. For further fitting processes the original model will be used.

Dataset	Model	A	B	C
0 - 50 dataset, 20°C data	Original	1.218±0.050	-1.859(±0.172) - 2.392(±0.127)i	1.963(±0.124) + 2.145(±0.109)i
	t fitting	1.218±0.050	-1.859(±0.416) - 2.392(±0.363)i	2.721(±0.097) + 1.025(±0.022)i
20 - 30 dataset, 20°C data	Original	0.362±0.024	-0.416(±0.086) - 0.665(±0.052)i	0.549(±0.056) + 0.485(±0.059)i
	t fitting	0.362±0.024	-0.416(±0.166) - 0.665(±0.134)i	0.632(±0.044) + 0.371(±0.016)i

TABLE 8.8: Non-linear fitted values obtained from original and t fitting for the 20°C data from both 75 mM PNP datasets

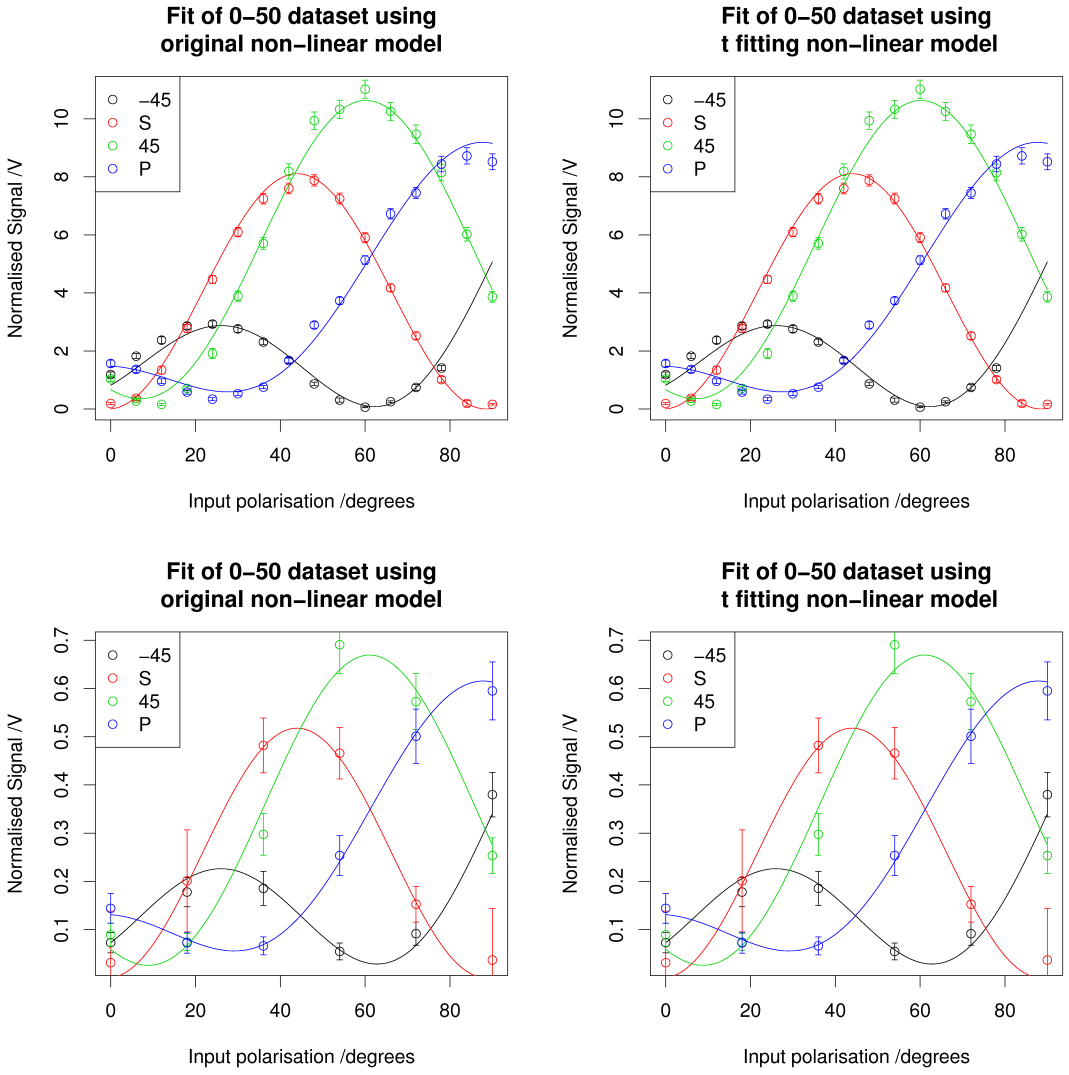


FIGURE 8.5: The summary plot showing the comparison between original and t fitting for the experiment at 20°C for both 75 mM PNP datasets. Error bars are $\pm 3\sigma$

8.3.2 Concentration dependence

The concentration dependence dataset is compiled of a number of experiments, shown in table 8.9. The surface coverages are determined using the fitting of the adsorption isotherms and are described in more detail below.

Conc. /mM	Approx surface coverage /%	Experiment number	Experimental data
75.12	88	597	S output P output 45 output -45 output
49.16	82	613	S output P output 45 output -45 output
29.92	74	599	S output P output 45 output -45 output
10.80	51	620	S output P output 45 output -45 output

TABLE 8.9: The combination of experiments used to generate the concentration dependence of surface orientation dataset for PNP at 20°C . The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

The results were fitted to Eq. 6.21 through an R script as with the temperature dependence data. Fig. 8.6 shows the fitted curves and table 8.10 shows the fitted coefficients.

Conc. /mM	A	B	C	Residual standard error
75.12	0.362±0.024	-0.416(±0.086) - 0.665(±0.052)i	0.549(±0.056) + 0.485(±0.059)i	0.027
49.16	0.390±0.027	-0.260(±0.099) - 0.706(±0.034)i	0.288(±0.059) + 0.548(±0.030)i	0.039
29.92	0.627±0.022	-0.784(±0.083) + 0.766(±0.081)i	0.803(±0.054) - 0.619(±0.066)i	0.042
10.80	0.719±0.016	-0.528(±0.058) + 0.627(±0.045)i	0.533(±0.031) - 0.308(±0.046)i	0.043

TABLE 8.10: Non-linear fitting values determined from the concentration dependence dataset at 20°C after applying a +3° offset to the output polarisation

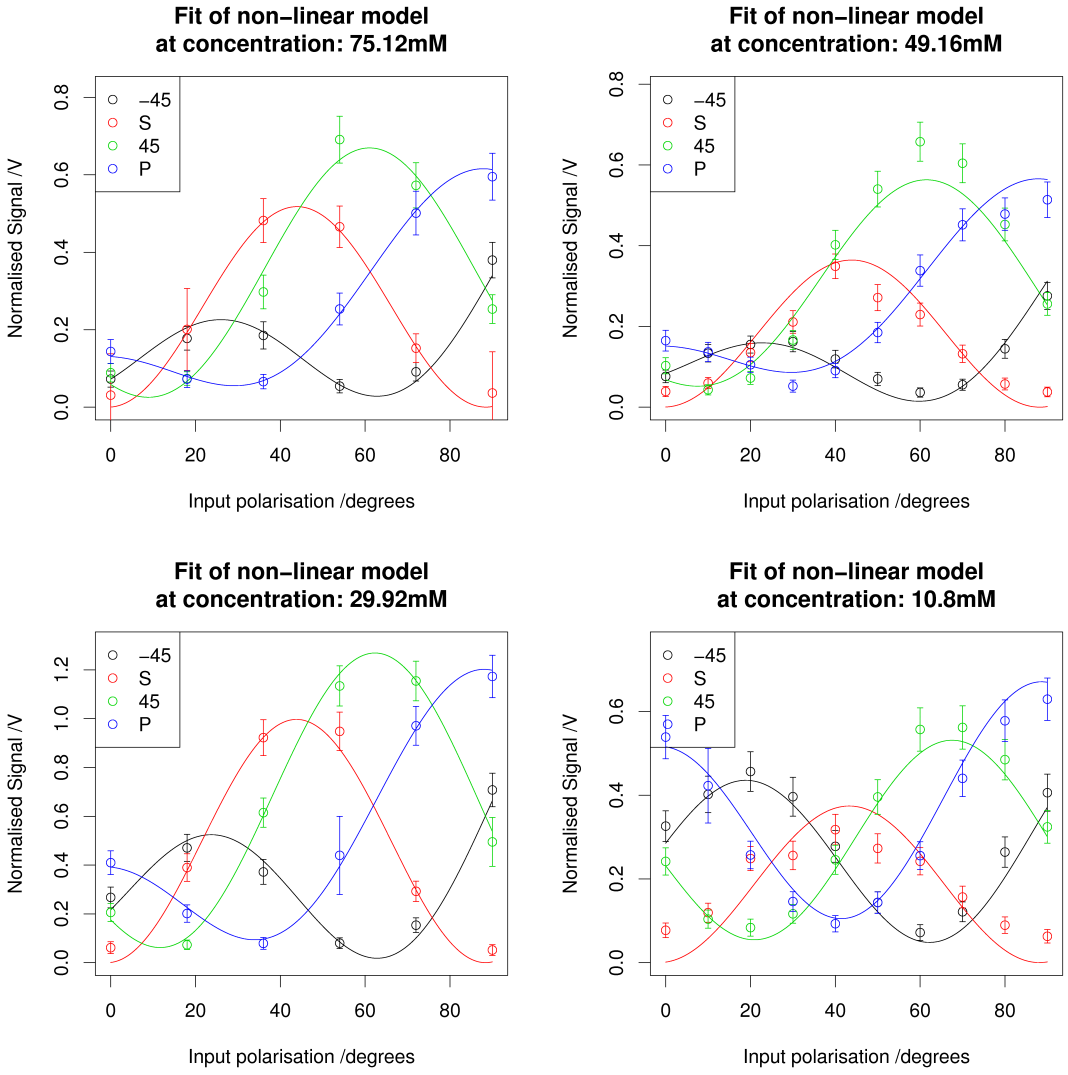


FIGURE 8.6: The summary plot of fitted curves against experimental data for each concentration from the concentration dependence dataset at 20^{o}C after applying the $+3^o$ offset to the output polarisation. Error bars are $\pm 3\sigma$

8.3.3 Adsorption isotherms

Each adsorption isotherm is made up from a number of experiments as with the previous studies. Table 8.11 describes how the dataset is made up.

Water bath temp. /^{o}C	Experiment number	Concentration range /mM
0	322	75 - 20.61
	323	20.61 - 5.64
	324	5.64 - 1.54
	324	0 - 3.82
10	351	74.76 - 20.46
	352	20.46 - 5.60
	353	5.64 - 1.53
	354	0 - 4.09
20	298	76.05 - 23.87
	299	23.87 - 7.49
	300	7.49 - 4.22
	303	0 - 4.06
30	357	20.49 - 5.61
	359	75.34 - 20.62
	332	0 - 4.06
40	385	107.80 - 45.95
	386	28.69 - 15.92
	387	14.15 - 4.90
	309	0 - 2.84
50	363	148.9 - 40.76
	365	25.45 - 17.87
	366	17.87 - 8.82
	367	8.82 - 4.35
	368	4.35 - 2.14
	369	0 - 2.02
	370	2.02 - 3.57
	371	3.57 - 4.74
	372	4.74 - 5.63

TABLE 8.11: The combination of experiments used to generate the concentration isotherm data for PNP. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

For each temperature and polarisation combination the data was fitted to Eq. 6.47 through an R script. The fitting used a non-linear regression to determine estimates of $\chi_{solvent}$, χ_{sample}, ϕ and the equilibrium constant, k. Fig. 8.7 shows the fitted curves plotted over the experimental data and table 8.12 gives the fitted coefficients.

Polarisations	Water bath temp. /^{o}C	$\chi_{solvent}$	χ_{sample}	cos(ϕ)	k /mol^{-1}dm^3
45/S	0	0.766±0.087	2.695±0.082	-0.992	267.099±20.514
	10	0.634±0.068	2.779±0.062	-0.965	135.825±8.429
	20	0.525±0.043	2.974±0.036	-0.943	94.786±3.246
	30	0.806±0.098	3.254±0.094	-0.997	113.403±9.765
	40	0.625±0.074	3.621±0.068	-0.868	42.061±2.495
	50	0.714±0.034	3.371±0.032	-0.868	57.534±1.859
54/45	0	0.978±0.073	3.217±0.068	-0.929	186.679±11.908
	20	0.822±0.036	3.839±0.033	-0.959	87.298±2.383
S/P	0	0.727±0.124	2.901±0.117	-1.000	271.620±26.506
	10	0.604±0.098	2.906±0.089	-1.000	153.753±13.167
	20	0.379±0.052	3.152±0.043	-1.000	94.782±4.137
	30	0.699±0.139	3.316±0.129	-1.000	119.888±12.477
	40	0.499±0.101	4.213±0.077	-0.943	32.421±2.623
	50	0.480±0.070	3.560±0.052	-0.721	44.757±1.776

TABLE 8.12: Non-linear fitting values determined from the concentration isotherm dataset

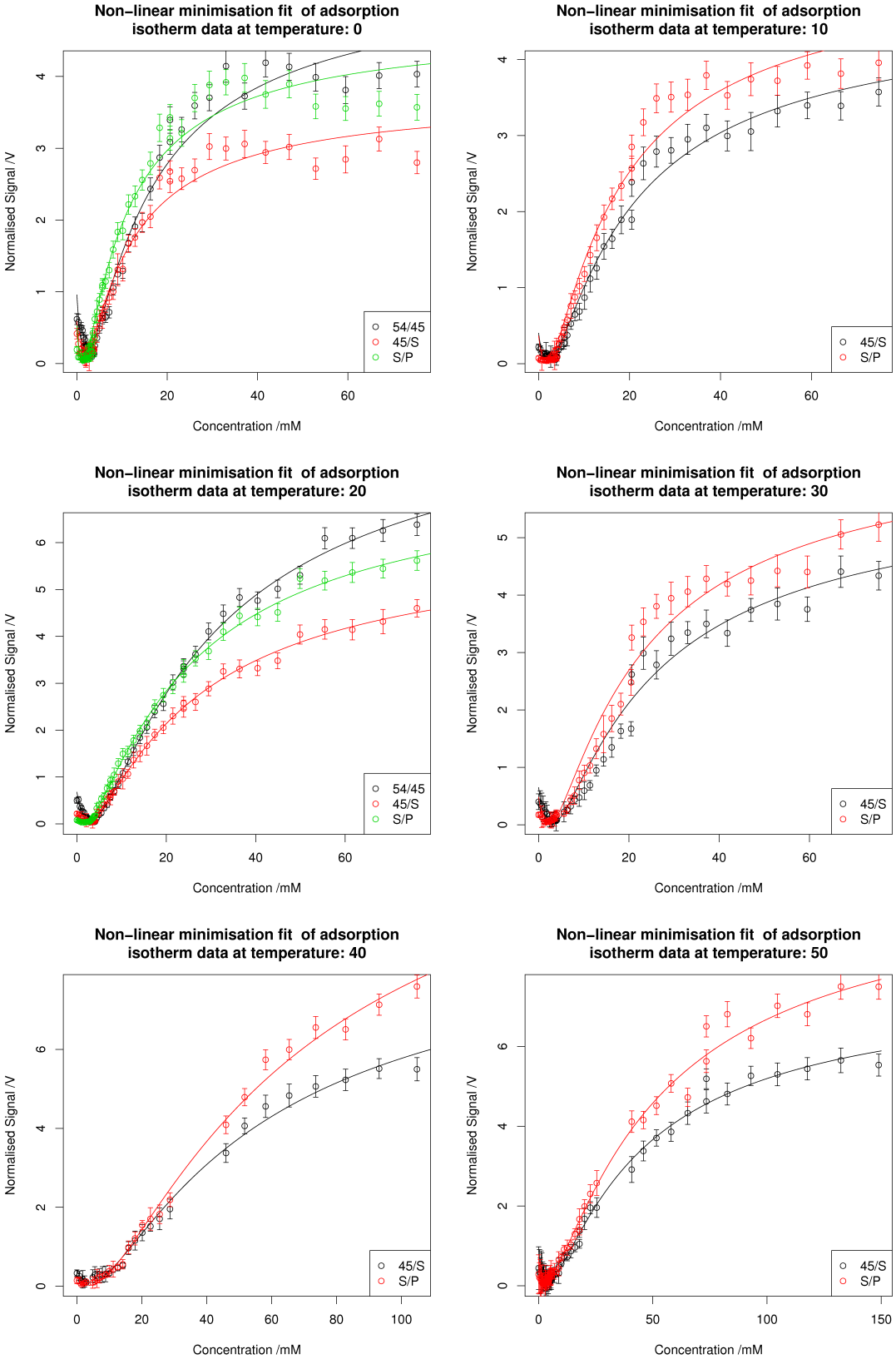


FIGURE 8.7: The summary plot of fitted curves against experimental data for each temperature from the adsorption isotherm dataset. Water bath temperatures are given. Error bars are $\pm 3\sigma$

8.4 Discussion

8.4.1 Surface orientation

8.4.1.1 Temperature dependence

As shown in Eq. 6.16, 6.17 and 6.18 the fitted values from the polarisation data relate to components of the electric susceptibility of the medium, χ. Even without determining the Fresnel coefficients, a_1-a_5, changes in the surface orientation can be observed by review of the ratios between these fitted coefficients.

Table 8.13 shows the ratio between B/A, C/A and B/C for the temperature dependence dataset (both 0-50 and 20-30 datasets have been included). The temperatures in this table have been converted to sample temperature as described in section 6.4.6. These results have been plotted and are shown in Fig. 8.8.

Dataset	Sample temp. /K	B/A	C/A	B/C
0 - 50	278.5	2.202±0.011	1.961±0.011	1.123±0.003
	286.1	2.379±0.021	2.191±0.019	1.086±0.004
	293.7	2.489±0.013	2.388±0.011	1.042±0.002
20 - 30	293.8	2.166±0.083	2.023±0.084	1.071±0.020
	295.2	2.183±0.127	2.170±0.123	1.006±0.027
	296.7	2.048±0.062	2.170±0.056	0.944±0.013
	298.3	2.600±0.120	2.379±0.124	1.093±0.021
	299.8	2.648±0.109	2.426±0.112	1.091±0.018
	301.4	2.560±0.087	2.386±0.090	1.073±0.015
0 - 50	301.3	2.879±0.021	2.672±0.020	1.078±0.003
	308.9	3.294±0.028	3.055±0.026	1.078±0.003
	316.6	3.921±0.053	3.619±0.049	1.083±0.004

TABLE 8.13: The ratio between fitted values from the 75 mM PNP temperature dependence datasets

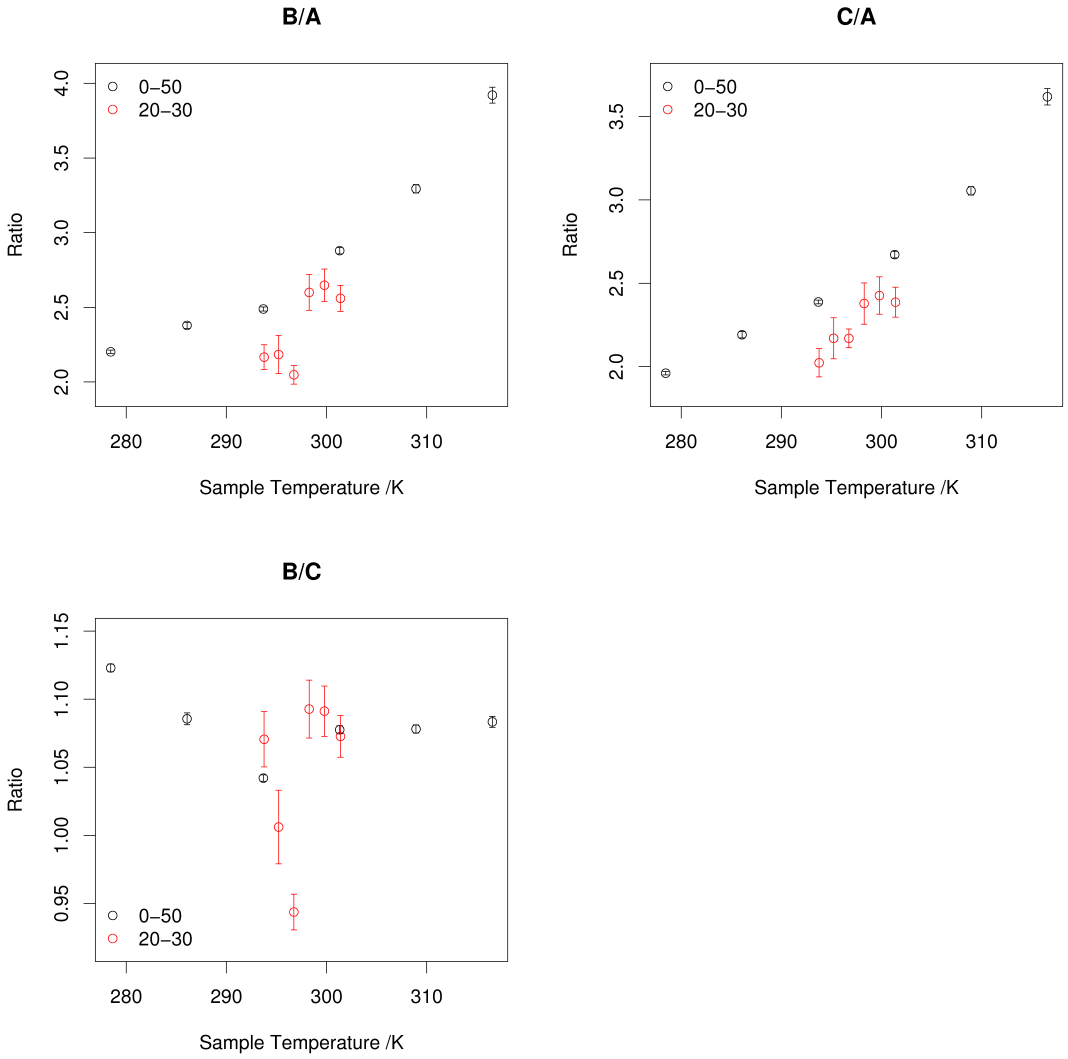


FIGURE 8.8: The plot of ratios between fitted values for the two 75 mM PNP datasets. Black indicates the 0-50 dataset and red the 20-30 dataset

These results show two distinct regions, the first running from 278 K to 293 K and the second from 293 K to 326 K. A trend line has been added to the plots in Fig. 8.8 to highlight this further and is shown in Fig. 8.9. The data suggests a phase transition occurring at approximately 296 K.

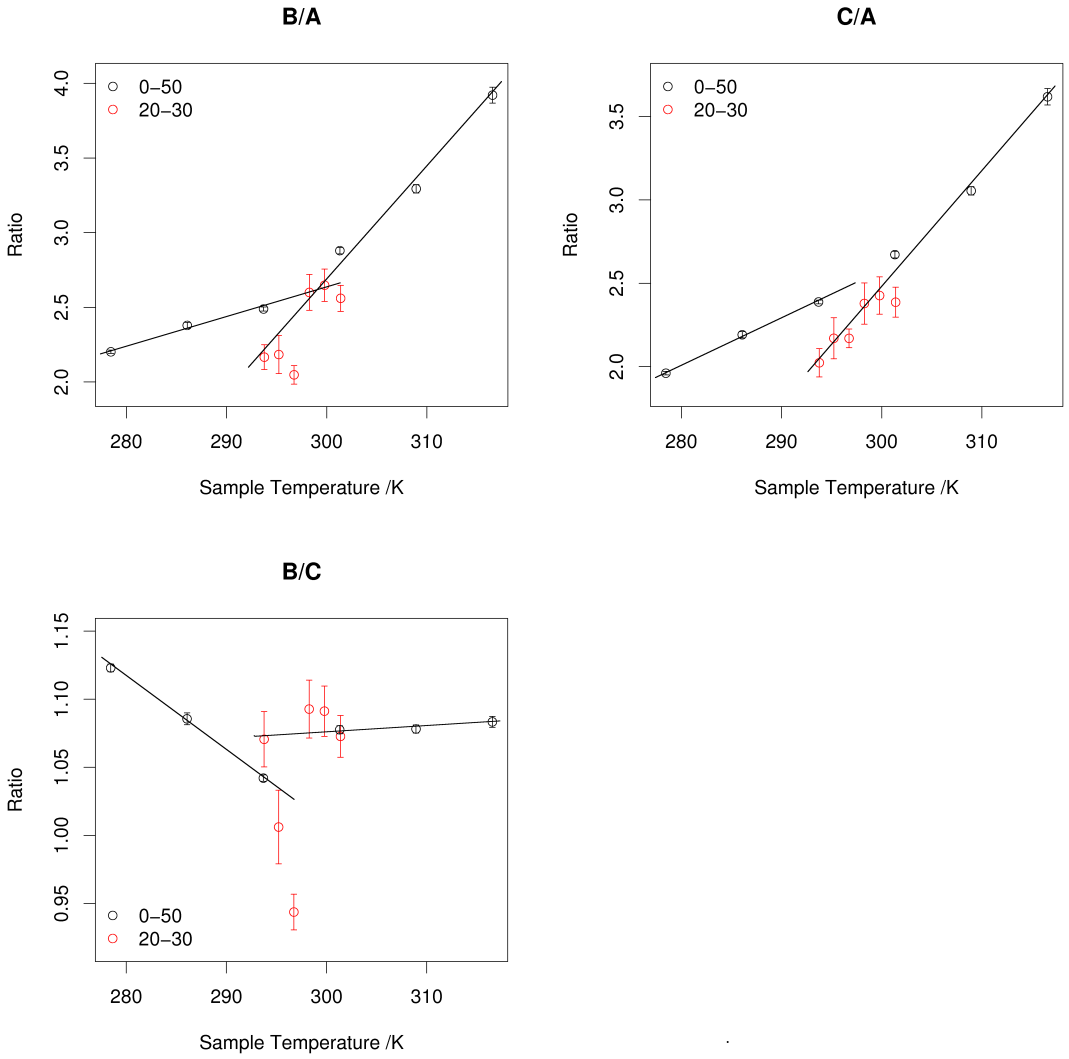


FIGURE 8.9: The plot of ratios between fitted values for the two 75 mM PNP datasets, including a simulated trend line. Black indicates the 0-50 dataset and red the 20-30 dataset

These experimentally determined parameters show information such as relative changes in the orientation at the surface but do not give specific information such as the molecular tilt angle, θ. To estimate these parameters, further assumptions must be made relating to the symmetry and orientation of the molecules at the surface, as described in section 6.1.2. Before these results can be submitted to Eq. 6.28 and 6.29/6.30 an estimate of the refractive index of the interface, n_3, must be obtained. In previous work this has been estimated as either the values of air or the bulk or the average between those values.

To obtain a better estimate of the refractive index at the interface a spectroscopic ellipsometry experiment was carried out. The instrument used in this experiment was a UVISEL VUV phase modulated spectroscopic ellipsometer from Horiba Jobin Yvon. A variable angle spectroscopic (VAS) experiment was completed to determine the optimal angle of incidence. The instrument was set up to run from 53^o to 54^o angle of incidence

in steps of 2^o . From this experiment, the optimal angle was found to be 53.3^o and as such the instrument set to this for further experiments. A spectroscopic experiment was carried out on a 75 mM PNP solution, scanning the wavelength from 300 nm to 700 nm in 2 nm steps. The results from this experiment were fitted as a Quatre Oscillator using a two layer model, the first oscillator relating to a dense layer of water at the surface[109] and the second relating to PNP. The water oscillator parameters were seeded with values determined in previous fitting and the PNP oscillator seeded with a strength corresponding to that seen in the UV-VIS spectrum. The model was configured to fit the thickness, angle of incidence, and oscillator parameters. The oscillator strength converged on that observed in the UV-VIS spectrum and the angle of incidence to that set for the experiment. The layer thickness was fitted as 2.22 ± 0.176 nm. Fig. 8.10 shows a plot of experimental results against fitted values. This fitting provided an estimate of the refractive index of the interface, n_3, at the fundamental and second harmonic wavelength as 1.405 and 1.574 respectively.

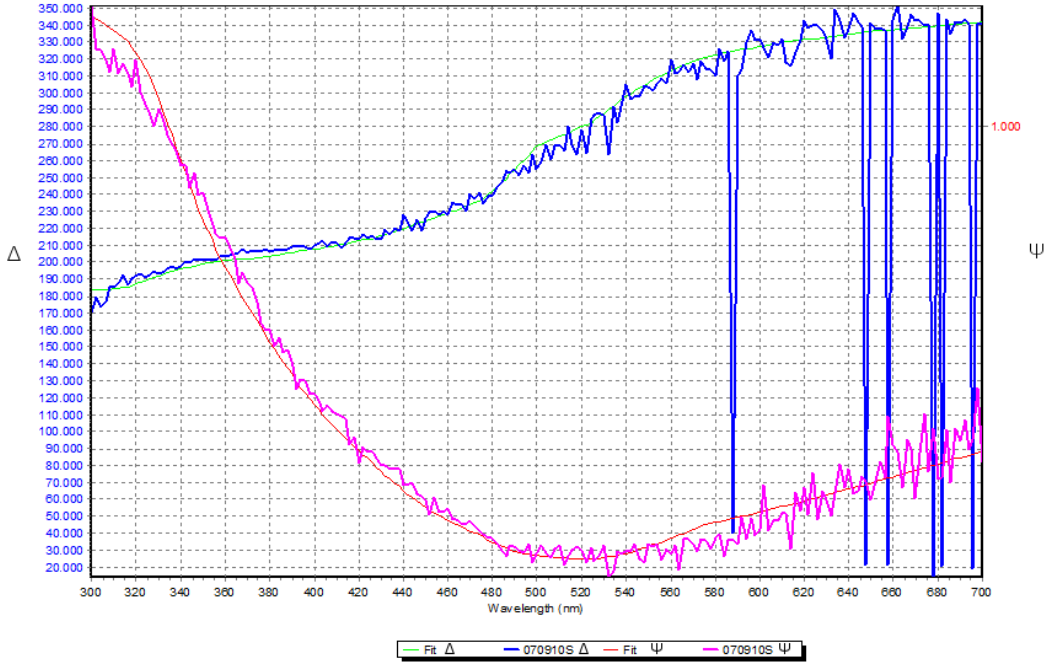


FIGURE 8.10: The plot of fitted curves against experimental results for the 75 mM PNP spectroscopic ellipsometry experiment

With these estimates of the refractive indices, the surface orientation parameter, D, and ratio between dominant hyperpolarisability terms were calculated and reported in table 8.14 and 8.15. The surface orientation parameter, D, has been reported for both a random ψ, rotation about the Z-axis, and fixed ψ of 90^o . Fig. 8.11 shows a plot of these parameters against sample temperature for the ψ fixed to 90^o results.

Dataset	Sample temp. /K	χ_{xxz}	χ_{zxx}	χ_{zzz}
0 - 50	278.45	9.303(±0.404) + 8.277(±0.426)i	-14.197(±1.020) - 18.527(±0.741)i	86.345(±2.739) + 62.143(±2.424)i
0 - 50	286.07	7.276(±0.615) + 8.404(±0.520)i	-12.873(±1.453) - 15.493(±1.142)i	70.989(±3.443) + 57.486(±2.812)i
0 - 50	293.69	-10.914(±0.687) - 11.925(±0.607)i	17.254(±1.599) + 22.204(±1.181)i	-99.342(±4.076) - 81.940(±3.371)i
20 - 30	293.69	3.053(±0.310) + 2.697(±0.326)i	-3.861(±0.802) - 6.171(±0.480)i	26.679(±2.238) + 20.476(±1.810)i
20 - 30	295.21	2.794(±0.382) + 2.793(±0.357)i	-2.428(±0.977) - 6.174(±0.369)i	22.446(±2.855) + 20.838(±1.789)i
20 - 30	296.74	3.8a56(±0.296) - 3.108(±0.346)i	-4.469(±0.715) + 6.395(±0.480)i	31.870(±2.045) - 22.386(±1.887)i
20 - 30	298.26	2.629(±0.740) - 4.360(±0.438)i	-4.751(±1.829) + 7.979(±1.033)i	27.161(±4.264) - 29.722(±2.377)i
20 - 30	299.79	2.378(±0.633) - 4.557(±0.320)i	-3.984(±1.627) + 8.473(±0.725)i	24.835(±3.939) - 31.293(±1.760)i
20 - 30	301.31	2.946(±0.443) - 4.094(±0.300)i	-3.594(±1.119) + 8.286(±0.464)i	26.266(±3.103) - 29.252(±1.635)i
0 - 50	301.31	10.571(±0.739) + 9.628(±0.770)i	-14.138(±1.740) - 21.484(±1.097)i	86.605(±4.925) + 72.175(±4.334)i
0 - 50	308.93	7.211(±1.203) - 14.506(±0.582)i	-12.095(±2.912) + 26.526(±1.254)i	70.435(±7.175) - 98.857(±3.217)i
0 - 50	316.55	13.641(±0.982) + 2.924(±4.346)i	-21.742(±3.848) - 12.800(±6.396)i	104.365(±9.069) + 32.536(±26.774)i

Table 8.14: The molecular susceptibilities determined from the two 75 mM PNP temperature dependence datasets

Dataset	Sample temp. /K	$\frac{\beta_{zxx}}{\beta_{zzz}}$	D ψ=90^o	D ψ random
0 - 50	278.45	-0.532(±0.003) - 0.112(±0.009)i	0.850(±0.001) - 0.006(±0.004)i	0.876(±0.001) - 0.001(±0.003)i
0 - 50	286.07	-0.545(±0.004) - 0.085(±0.018)i	0.847(±0.002) - 0.017(±0.007)i	0.874(±0.001) - 0.011(±0.007)i
0 - 50	293.69	-0.543(±0.003) - 0.089(±0.013)i	0.842(±0.001) - 0.012(±0.005)i	0.870(±0.001) - 0.007(±0.005)i
20 - 30	293.69	-0.523(±0.007) - 0.128(±0.020)i	0.846(±0.003) - 0.001(±0.007)i	0.872(±0.002) + 0.004(±0.006)i
20 - 30	295.21	-0.517(±0.009) - 0.153(±0.023)i	0.837(±0.004) + 0.005(±0.008)i	0.866(±0.003) + 0.010(±0.007)i
20 - 30	296.74	-0.504(±0.005) + 0.116(±0.016)i	0.838(±0.002) + 0.001(±0.006)i	0.866(±0.002) - 0.003(±0.005)i
20 - 30	298.26	-0.564(±0.012) + 0.092(±0.045)i	0.843(±0.005) + 0.019(±0.019)i	0.871(±0.004) + 0.013(±0.017)i
20 - 30	299.79	-0.570(±0.011) + 0.102(±0.037)i	0.841(±0.004) + 0.018(±0.015)i	0.870(±0.003) + 0.011(±0.013)i
20 - 30	301.31	-0.553(±0.008) + 0.126(±0.022)i	0.840(±0.003) + 0.006(±0.008)i	0.869(±0.002) + 0.001(±0.007)i
0 - 50	301.31	-0.552(±0.004) - 0.109(±0.011)i	0.842(±0.002) + 0.001(±0.005)i	0.870(±0.001) + 0.005(±0.004)i
0 - 50	308.93	-0.585(±0.006) + 0.086(±0.021)i	0.837(±0.002) + 0.015(±0.009)i	0.867(±0.002) + 0.009(±0.008)i
0 - 50	316.55	-0.560(±0.006) - 0.076(±0.022)i	0.842(±0.002) + 0.016(±0.008)i	0.871(±0.001) + 0.016(±0.007)i

TABLE 8.15: The molecular orientation parameters determined from the two 75 mM PNP temperature dependence datasets

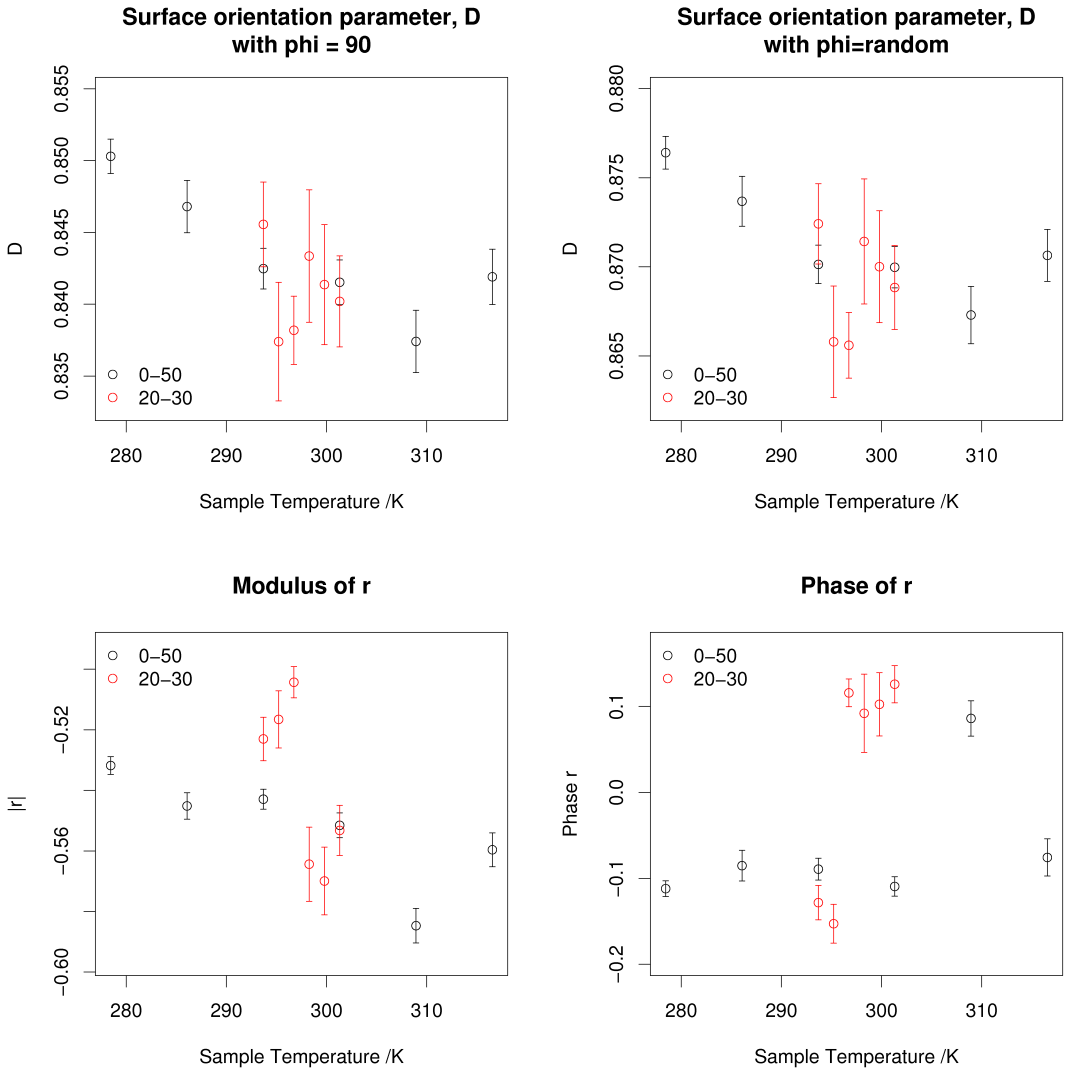


FIGURE 8.11: The plot of surface orientation parameter, D, and ratio between dominant hyperpolarisability terms against sample temperature for the 75 mM PNP temperature dependence datasets

This plot suggests the trends seen in the experimentally determined results, seen in Fig. 8.9, are carried through to the molecular properties. The plots of the orientation parameter, D, showing a steady change between 0 and 20^{o}C and with error remaining constant above 30^{o}C . The data between 20 and 30^{o}C deviates from these trends somewhat, this may be due to the shorter experimental procedure carried out in this dataset leading to less well determined parameters. The changes observed in molecule tilt angle correspond to an average molecular shift of 1^o across the temperature range.

These results show little complex component in the D components in either the random or fixed ψ models, suggesting either model is applicable. This has been investigated further using the F, D and R fitting described in chapter 6. The experimental data was fitted via non-linear regression to Eq. 6.40 allowing F to be fitted while fixing D and R to that determined through the previous fitting. The results from this process are

shown in table 8.16. Fig. 8.12 showing a summary of the fitting for random ψ and Fig. 8.13 showing a summary for ψ fixed to 90^o .

ψ assumption	Water bath temp. ^{o} C	F	D	R	RSE
90degrees	0	65.620±0.236	0.850	-0.532 - 0.112i	0.181
	10	56.818±0.297	0.847	-0.545 - 0.085i	0.204
	20	80.591±0.320	0.842	-0.543 - 0.089i	0.311
	20	20.881±0.166	0.846	-0.523 - 0.128i	0.025
	22	19.264±0.201	0.837	-0.517 - 0.153i	0.029
	24	24.418±0.155	0.838	-0.504 + 0.116i	0.026
	26	25.211±0.285	0.843	-0.564 + 0.092i	0.057
	28	25.098±0.254	0.841	-0.570 + 0.102i	0.052
	30	24.683±0.188	0.840	-0.553 + 0.126i	0.037
	30	70.655±0.289	0.842	-0.552 - 0.109i	0.262
	40	76.901±0.383	0.837	-0.585 + 0.086i	0.419
	50	68.740±0.375	0.842	-0.560 - 0.076i	0.330
Random	0	66.135±1.006	0.876	-0.532 - 0.112i	0.746
	10	57.470±0.954	0.874	-0.545 - 0.085i	0.627
	20	82.489±1.354	0.870	-0.543 - 0.089i	1.246
	20	20.910±0.543	0.872	-0.523 - 0.128i	0.078
	22	19.497±0.525	0.866	-0.517 - 0.153i	0.072
	24	24.804±0.670	0.866	-0.504 + 0.116i	0.106
	26	25.194±0.726	0.871	-0.564 + 0.092i	0.140
	28	25.110±0.714	0.870	-0.570 + 0.102i	0.142
	30	24.785±0.663	0.869	-0.553 + 0.126i	0.126
	30	71.921±1.125	0.870	-0.552 - 0.109i	0.973
	40	78.243±1.318	0.867	-0.585 + 0.086i	1.373
	50	70.014±1.121	0.871	-0.560 - 0.076i	0.941

TABLE 8.16: The experimentally determined (D and R) and fitted (F) value from the two 75 mM PNP temperature dependence datasets fitted to the F, D and R model. (RSE - Residual standard error)

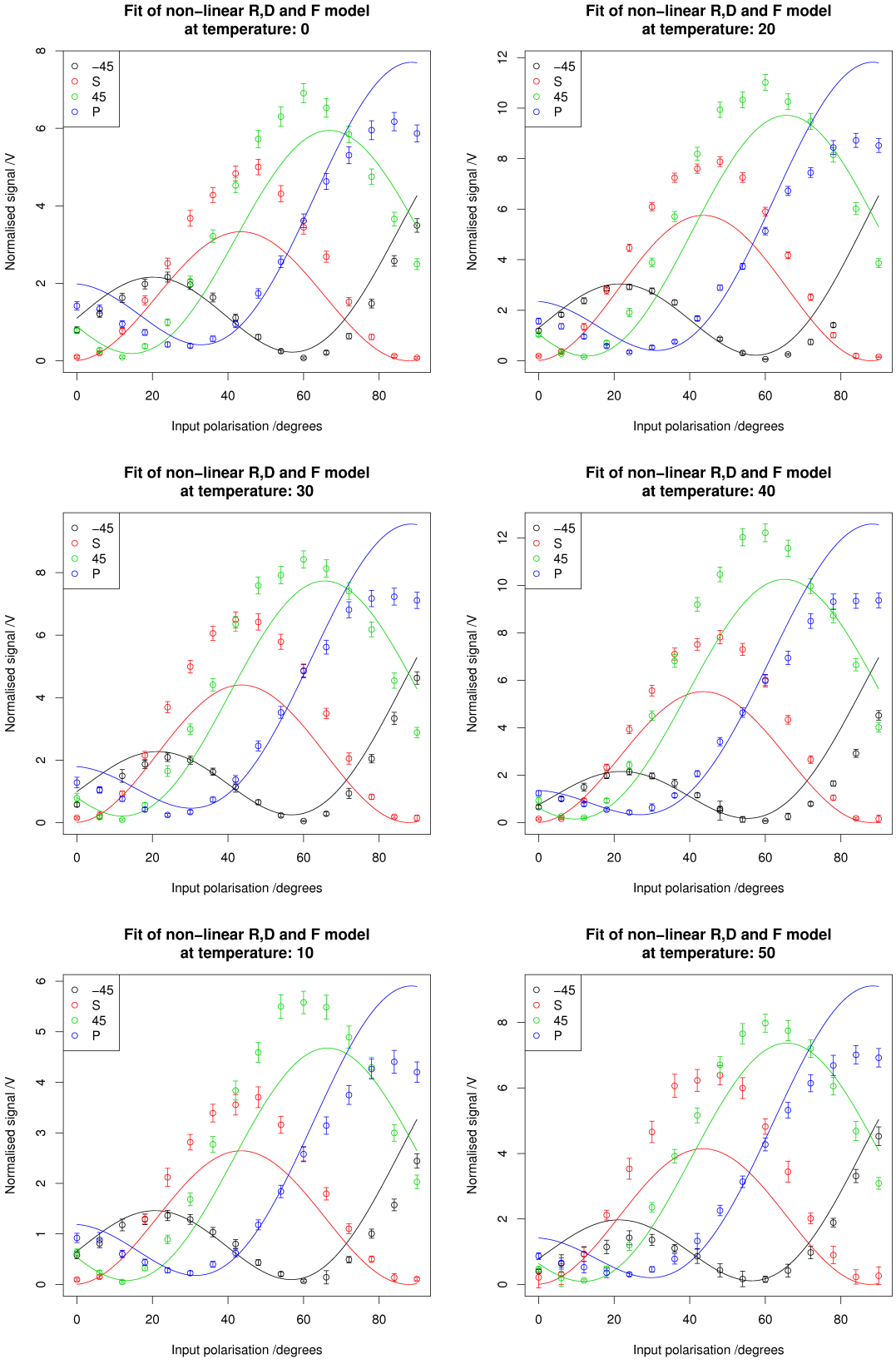


FIGURE 8.12: The summary plot of fitted curves against experimental data for each temperature from the 0-50^{o}C 75 mM PNP dataset fitted to the model assuming random ψ. Water bath temperatures are given

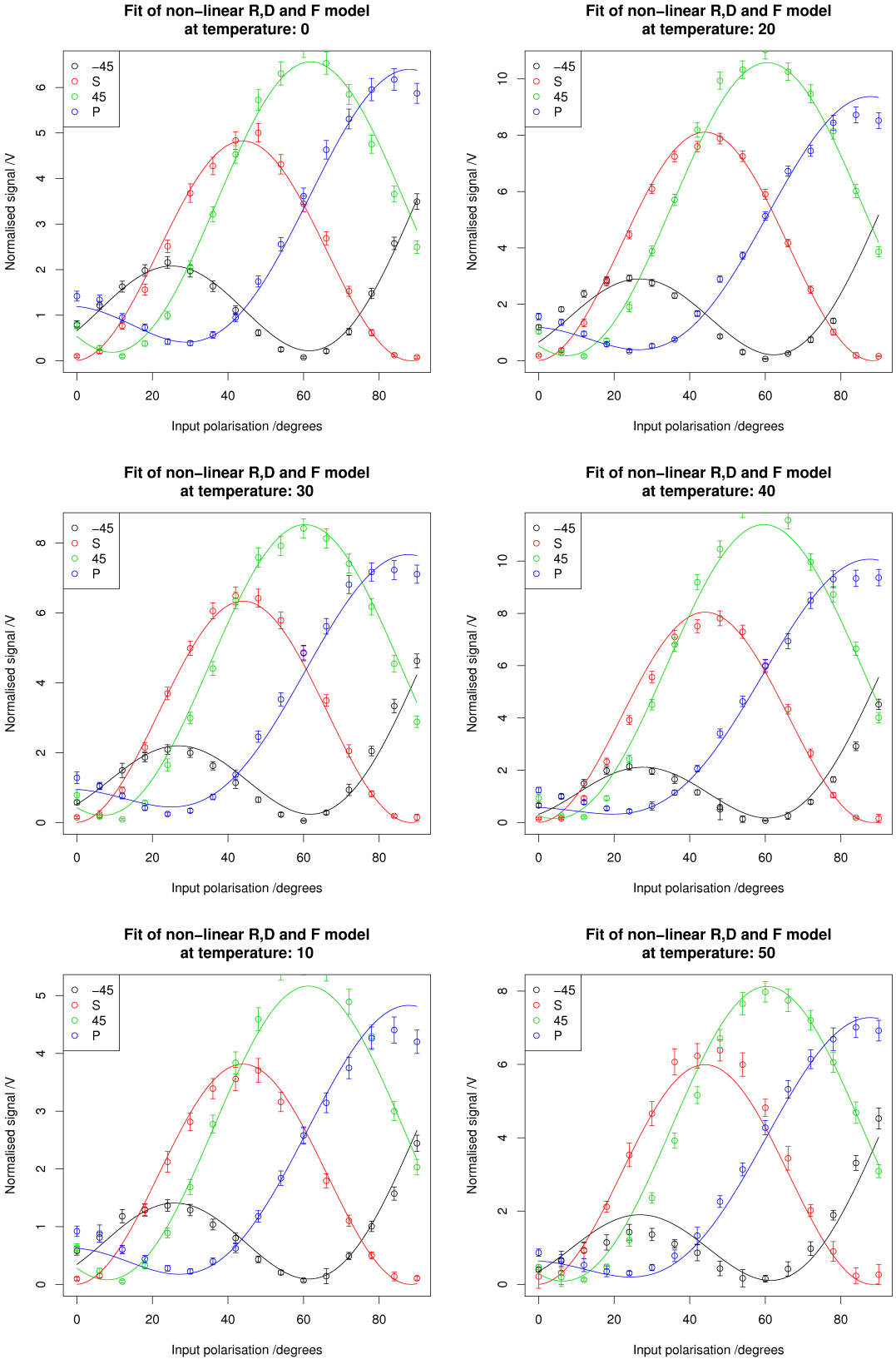


FIGURE 8.13: The summary plot of fitted curves against experimental data for each temperature from the 0-50^{o}C 75 mM PNP dataset fitted to the model assuming ψ fixed to 90^o . Water bath temperatures are given

It can be seen from these results there is a significant difference in the fits for the two models. The results from ψ fixed to 90^o show reduced error in the regression results at each temperature and show a good fit in the plots, compared to the large deviation from experimental results when ψ is considered random. This shows that although both models can be considered applicable when fitting via A, B and C due to the small imaginary component in D, only fixing ψ to 90^o will give accurate results for PNP at the air/liquid interface.

The molecular tilt angle determined in this work differs from that calculated in previous studies, shown in table 8.17. It is believed this is due to the additional accuracy and repeatability of the experimental process as well as the new techniques used to more accurately estimate the refractive indices at fundamental and second harmonic wavelengths. These experiments were also carried out at differing PNP concentration which, as discussed below, will have an effect on the molecular tilt angle.

Source	PNP concentration /mM	Orientation parameter, D
This work	75	0.844 ± 0.001
Timson *et al.*[104]	7	0.61
Bell *et al.*[110]	100	0.63
Tamburello-Luca *et al.*[115]	50	0.26
Sarker *et al.*[122]	50	0.60

TABLE 8.17: The experimentally determined values of the orientation parameter, D, for PNP across multiple studies

These results have determined an average molecular tilt angle, θ, of 23^o . This is the angle between the surface normal and the primary molecular angle, in PNP this is the C_2 axis; shown in Fig. 8.14

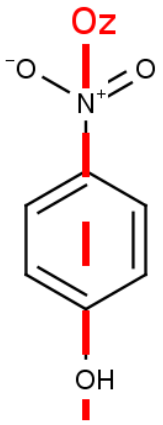


FIGURE 8.14: The primary axis of PNP

The R, D and F fitting shows that the rotation about this primary axis, ψ, is also fixed relative to the other molecules. This suggests there are π interactions between neighbouring molecules producing a favourable orientation. Fig. 8.15 shows a representation of how these angles are translated to molecules at the surface.

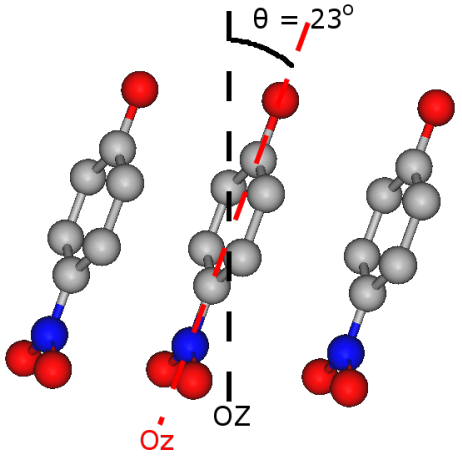


FIGURE 8.15: A representation of the alignment of the molecules at the surface based on the angles determined through this study

The results determined through this work cannot calculate the absolute phase of the determined parameters and therefore cannot yield which group will point towards either medium. This can be determined through comparison of parameters determined for the sample to that of an known reference such as a z-cut quartz crystal.

8.4.1.2 Concentration dependence

The ratio between fitted values have also been reported for the concentration dependence experiments. Table 8.18 shows the ratio between the fitted values at each concentration and Fig. 8.16 shows the plots of this data.

Concentration /mM	B/A	C/A	B/C
75.12	2.166±0.083	2.023±0.084	1.071±0.020
49.16	1.930±0.087	1.588±0.091	1.215±0.035
29.92	1.749±0.030	1.618±0.031	1.080±0.012
10.80	1.140±0.027	0.856±0.029	1.332±0.036

TABLE 8.18: The ratio between fitted values from the PNP concentration dependence results at 20^{o}C

Using Eq. 6.45 the concentration can be converted to surface coverage via the equilibrium constant, k. Using the fitting results from the adsorption isotherms at 20^{o}C ,

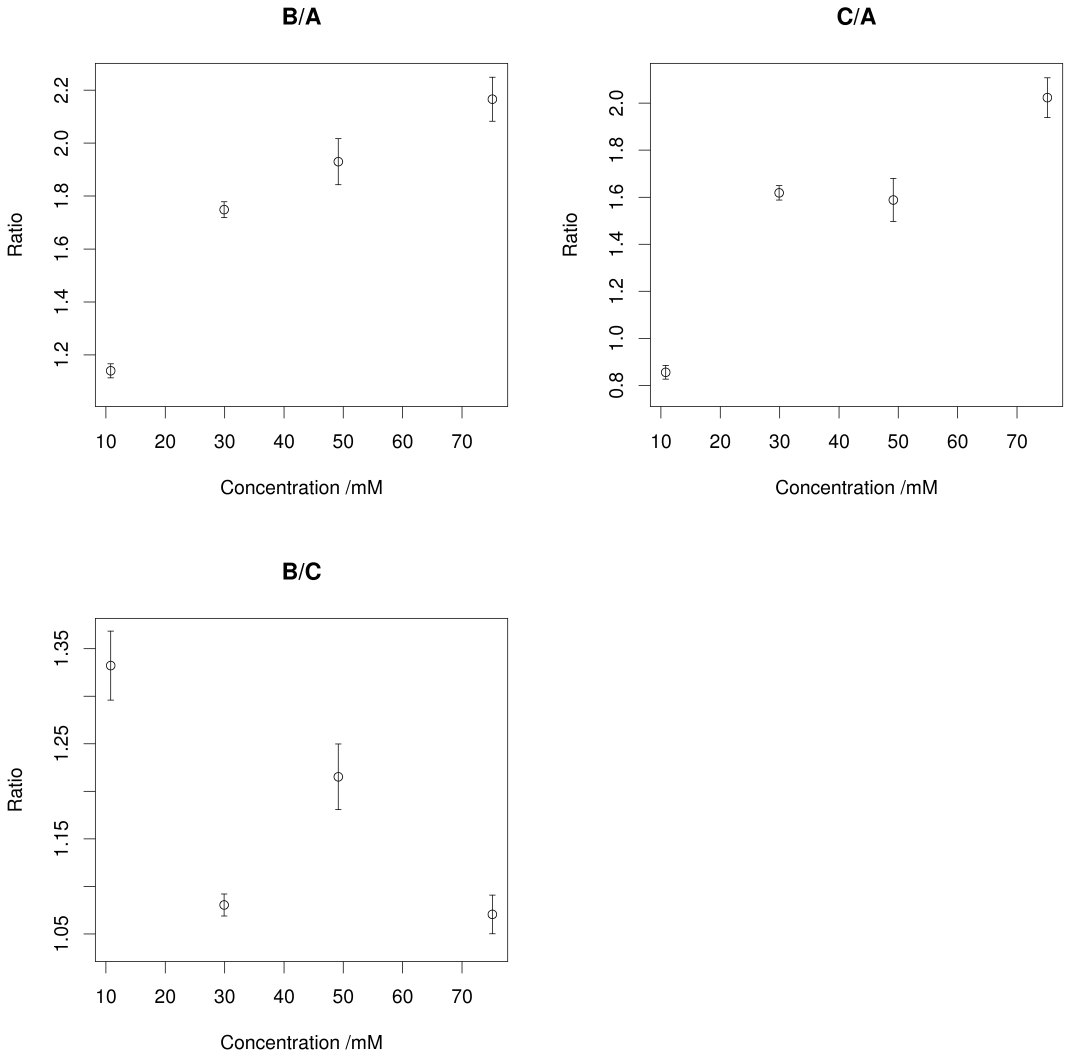


FIGURE 8.16: The plot of ratios between fitted values against concentration for the PNP concentration dependence results at 20^{o}C

discussed in section 8.3.3, we can calculate the coverage at each concentration. A plot showing the ratio between fitted values and coverage is shown in Fig. 8.17.

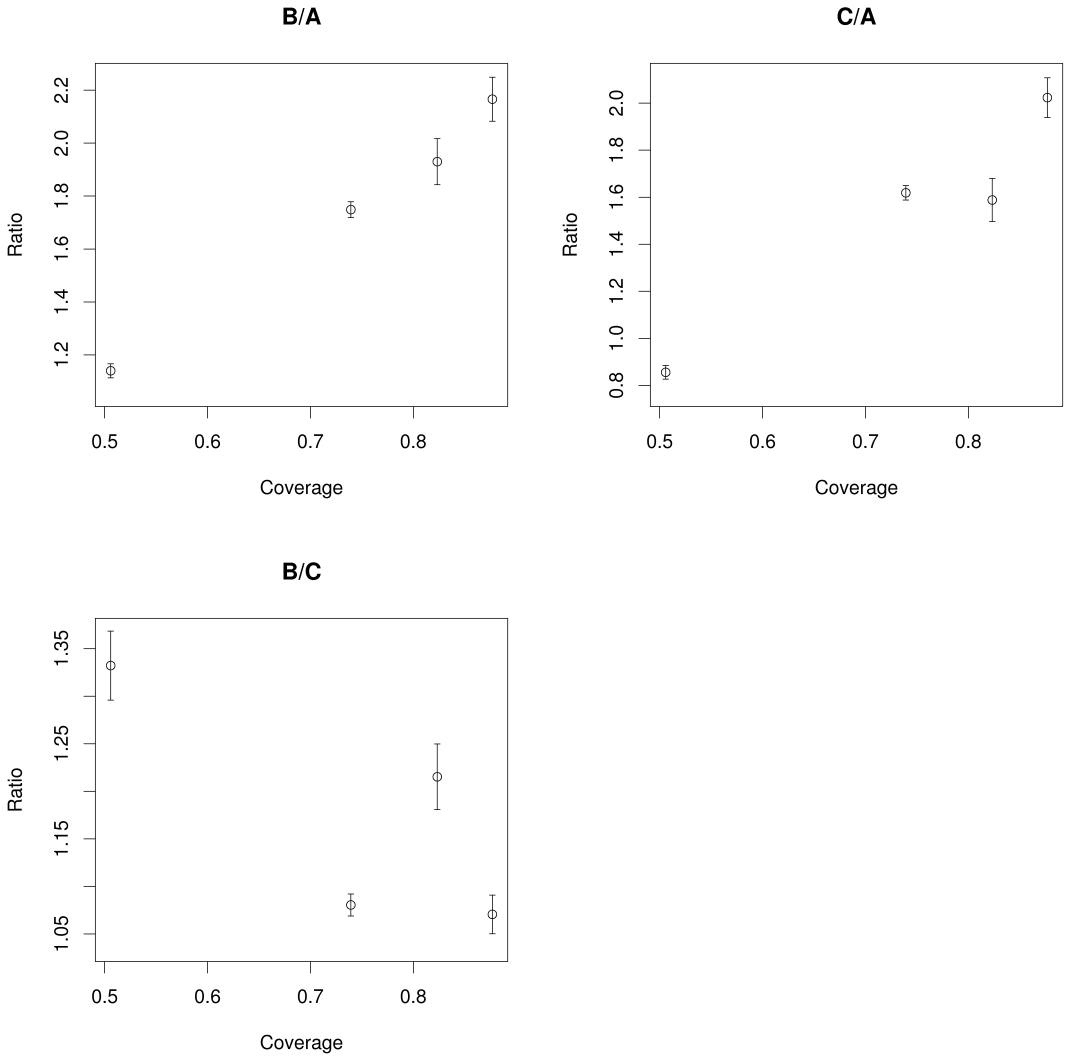


FIGURE 8.17: The plot of ratios between fitted values against coverage for the PNP concentration dependence results at 20^{o}C

From this plot we can see the surface parameters change linearly with coverage. This suggests that as more compounds are present at the surface, due to the higher bulk concentration, the molecules re-align to minimise unfavourable interactions. This will continue until the point where the surface is full; at this point the system may change to include multiple layers or simply not allow further molecules at the surface.

As with the temperature dependence datasets, these results were processed further to calculate the surface orientation parameter, D, and ratio between dominant hyperpolarisability parameters. The results were fitted to Eqs. 6.28 and 6.29/6.30 using the refractive indices at the interface as determined through the ellipsometry experiment described above. These values are described in table 8.19 and 8.20. Fig. 8.18 shows a plot of these results against concentration and Fig. 8.19 against coverage.

Conc. /mM	Coverage /%	χ_{xxz}	χ_{zxx} χ_{xxz}	χ_{zzz}
75.12	88	3.053($\pm$0.310) + 2.697($\pm$0.326)i	-3.861($\pm$0.802) - 6.171($\pm$0.480)i	26.679($\pm$2.238) + 20.476($\pm$1.810)i
49.16	82	1.602($\pm$0.325) + 3.045($\pm$0.165)i	-2.409($\pm$0.922) - 6.550($\pm$0.316)i	19.491($\pm$2.370) + 22.413($\pm$0.853)i
29.92	75	4.467($\pm$0.299) - 3.444($\pm$0.368)i	-7.279($\pm$0.767) + 7.107($\pm$0.749)i	44.146($\pm$1.957) - 24.837($\pm$2.090)i
10.80	51	2.961($\pm$0.172) - 1.712($\pm$0.256)i	-4.901($\pm$0.539) + 5.815($\pm$0.418)i	36.748($\pm$1.410) - 16.208($\pm$1.371)i

TABLE 8.19: The molecular susceptibilities determined from the PNP concentration dependence datasets completed at 20°C

Conc. /mM	Coverage /%	$\frac{\beta_{zxx}}{\beta_{zzz}}$	D ϕ=90°	D ϕ random
75.12	88	-0.523(±0.007) - 0.128(±0.020)i	0.846(±0.003) - 0.001(±0.007)i	0.872(±0.002) + 0.004(±0.006)i
49.16	82	-0.549(±0.011) - 0.156(±0.027)i	0.853(±0.004) - 0.019(±0.010)i	0.879(±0.003) - 0.011(±0.009)i
29.92	74	-0.501(±0.005) + 0.098(±0.019)i	0.855(±0.002) + 0.011(±0.007)i	0.879(±0.002) + 0.006(±0.006)i
10.80	51	-0.438(±0.006) + 0.151(±0.017)i	0.881(±0.003) + 0.004(±0.006)i	0.899(±0.002) - 0.001(±0.005)i

TABLE 8.20: The molecular orientation parameters determined from the PNP concentration dependence datasets completed at 20°C

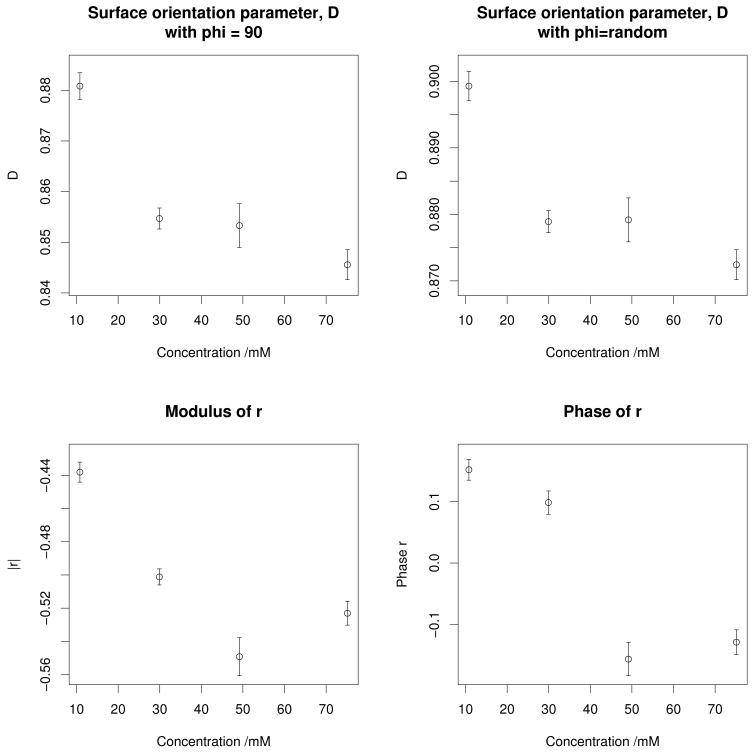


FIGURE 8.18: The plot of surface orientation parameter, D, and ratio between dominant hyperpolarisability terms against concentration for the PNP concentration dependence datasets run at 20^oC

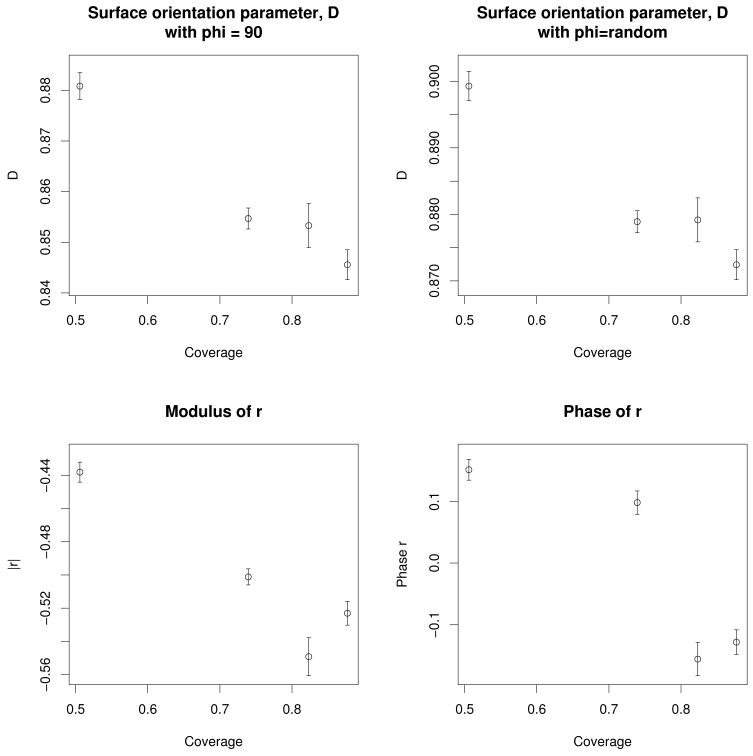


FIGURE 8.19: The plot of the surface orientation parameter, D, and ratio between dominant hyperpolarisability terms against coverage for the PNP concentration dependence datasets run at 20^oC

These results mirror those seen in the ratios between experimentally determined parameters, A, B and C. The plots show a linear trend in surface orientation parameter, D, and real components of the hyperpolarisability with coverage. The results show the average molecular tilt angle at the lower coverage, 50%, at an angle of 20^o which shifts to 23^o at a coverage of 88%.

8.4.2 Adsorption isotherms

The results from the adsorption isotherms were fitted to Eq. 6.47, which assumes the surface follows the Langmuir adsorption model. The results show a good fit to the data. The 0^oC results show some variation from the fitted curve at the higher concentrations. As the experiment used serial dilution to get to the lower concentrations these results would have been the first to be collected. After further research into PNP at the air/liquid interface it has been shown there is a settling time of the surface, which will be discussed further in chapter 11. It is thought the deviate from the fitted curve at high concentrations due to this settling effect.

From the fitted values, the adsorption free energy can be calculated from Eq. 8.5.

$$\Delta G^{\ominus}_{ads} = -RTln(K) \tag{8.5}$$

where R is the Gas constant, T is the temperature of the sample and K=55.5k. k is the fitted equilibrium constant and 55.5 is the molarity of water[118].

Table 8.21 shows the equilibrium constant, k, and adsorption free energy for each temperature and polarisation combination. Fig. 8.20 shows a plot of this data. The temperatures in these results have been converted to sample temperature, as described in section 6.4.6.

Polarisations	Sample Temperature /K	Equilibrium constant, k /mol^{-1}dm^3	Ln(K)	Adsorption free energy, Δ_{ads}G /kJmol^{-1}
45/S	278.5	267.099±20.514	9.604±0.077	-22.235±0.178
	286.1	135.825±8.429	8.928±0.062	-21.235±0.148
	293.7	94.786±3.246	8.568±0.034	-20.922±0.084
	301.3	113.403±9.765	8.747±0.086	-21.914±0.216
	308.9	42.061±2.495	7.756±0.059	-19.921±0.152
	316.6	57.534±1.859	8.069±0.032	-21.237±0.085
54/-45	278.5	186.679±11.908	9.246±0.064	-21.406±0.148
	293.7	87.298±2.383	8.486±0.027	-20.721±0.067
S/P	278.5	271.620±26.506	9.621±0.098	-22.274±0.226
	286.1	153.753±13.167	9.052±0.086	-21.530±0.204
	293.7	94.782±4.137	8.568±0.044	-20.922±0.107
	301.3	119.888±12.477	8.803±0.104	-22.053±0.261
	308.9	32.421±2.623	7.495±0.081	-19.252±0.208
	316.6	44.757±1.776	7.818±0.040	-20.576±0.104

TABLE 8.21: The calculated surface adsorption free energy from the PNP adsorption isotherm dataset

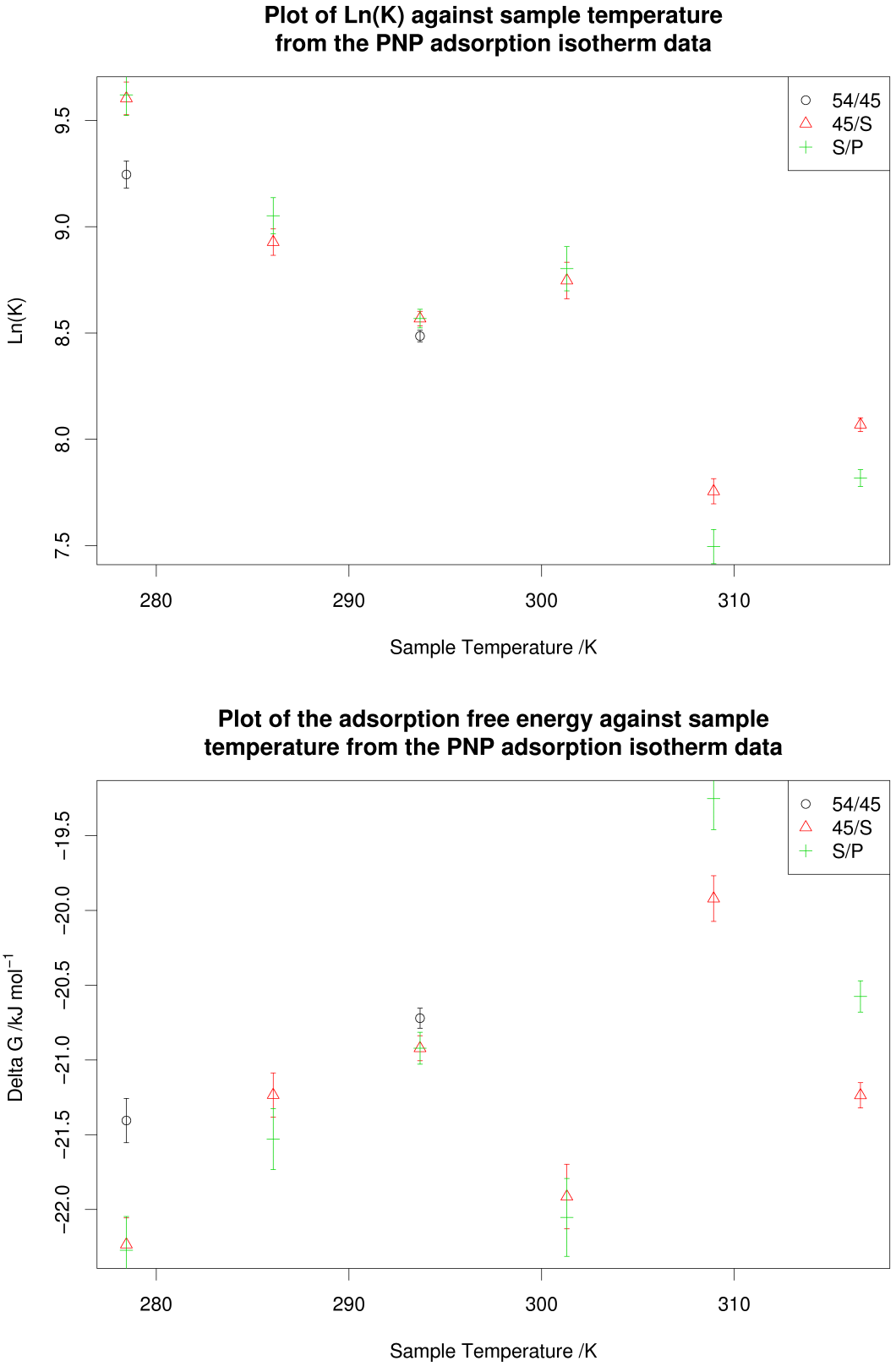
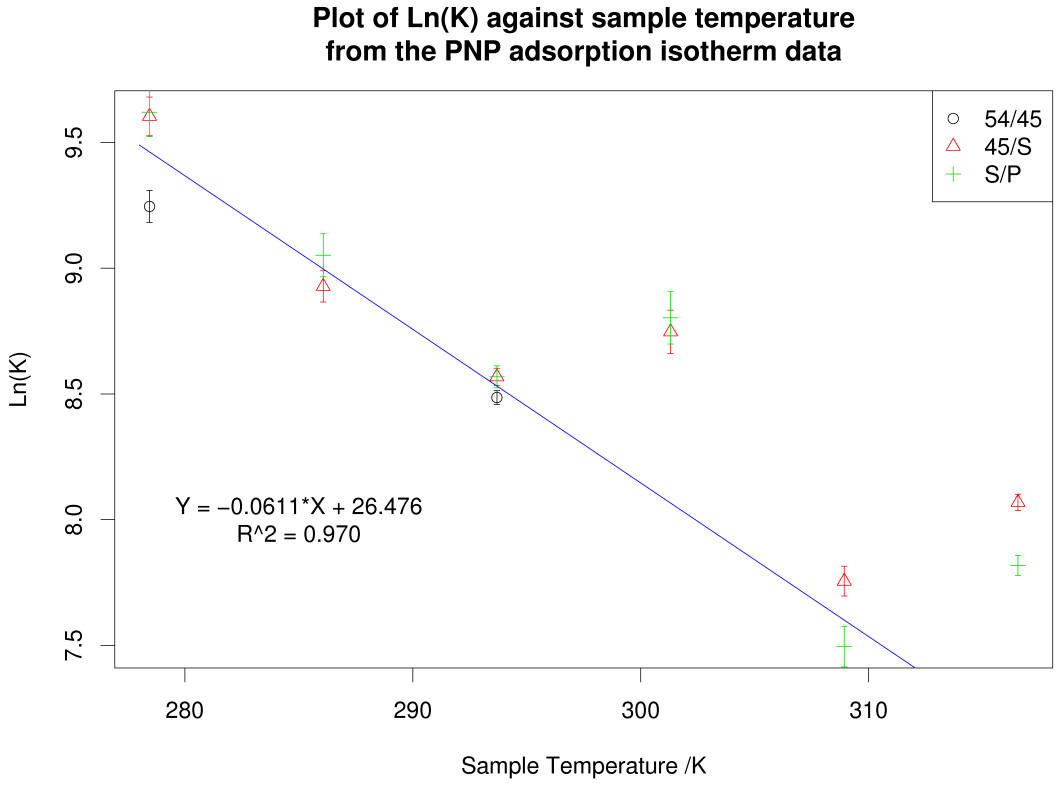


FIGURE 8.20: A plot of Ln(equilibrium constant) and surface adsorption free energy from the PNP adsorption isotherm dataset

If the 30^{o}C and 50^{o}C data is considered outliers then the plot shows a linear trend in the surface adsorption free energy with temperature; this is shown in Fig. 8.21. When reviewing the raw data, it can be seen the 30^{o}C data does not show a smooth change in signal with concentration as with the other data. In a standard adsorption isotherm a single solution underwent a number of serial solutions from the highest concentration to the lowest in one run, covering a number of experiments. Due to issues in running the experiment this was not possible for the 30^{o}C isotherm and as such the data is made up from two solutions, the first from 75 mM to 20 mM and the second from 20 mM to 5 mM. This change in solution between experiments may account for the deviation from the expected results.

From these results we can determine the enthalpy, $\Delta_{ads}H$, and entropy, $\Delta_{ads}S$, of the system at 298K as -43.63 $\pm$ 2.64 kJ mol^{-1} and 77.7 $\pm$ 9.11 J mol^{-1} K^{-1} respectively. When comparing the adsorption free energy calculated in this study, -20.46 kJ mol^{-1} at 298K, to that in a previous study[116], -19 kJ mol^{-1} (assumed 298K), it can be seen the values are similar. The values determined in this study are calculated from a number of isotherms at multiple temperatures, compared to a single isotherm in the previous study. The results from this work should therefore provide more accurate results. This experimental method has shown a novel approach to the determination of entropy of adsorption to the surface.

These results can also be compared to those determined by Paluch *et al.*[114] and re-evaluated by Tamburello-Luca *et al.*[115]. These experiments used the maximum bubble pressure method to measure the surface tension of PNP over range of concentrations and temperatures. These results were fitted to a Volmer's isotherm and the Gibbs free energy of adsorption to the surface was determined; from this the enthalpy and entropy of adsorption were also calculated. Tamburello-Luca *et al.* re-fitted the experimental data using a Frumkin isotherm to determine a second estimate of the Gibbs free energy of adsorption to the surface. Paluch *et at.* calculated the Gibbs free energy of adsorption to the surface as -10.51 kJ mol^{-1}, the enthalpy of adsorption to the surface as -8.37 kJ mol^{-1} and the entropy of adsorption to the surface as 7.28 J mol-1 K^{-1}. Using a Frumkin fit, Tamburello-Luca *et al.* re-calculated the Gibbs free energy of adsorption to the surface as -16.4 $\pm$ 0.2 kJ mol^{-1}. The variation of these results to those calculated in this study may come from either the experimental method or the analytical process. It has been already been shown in the variation between the determined parameter how significant the analytical method is.



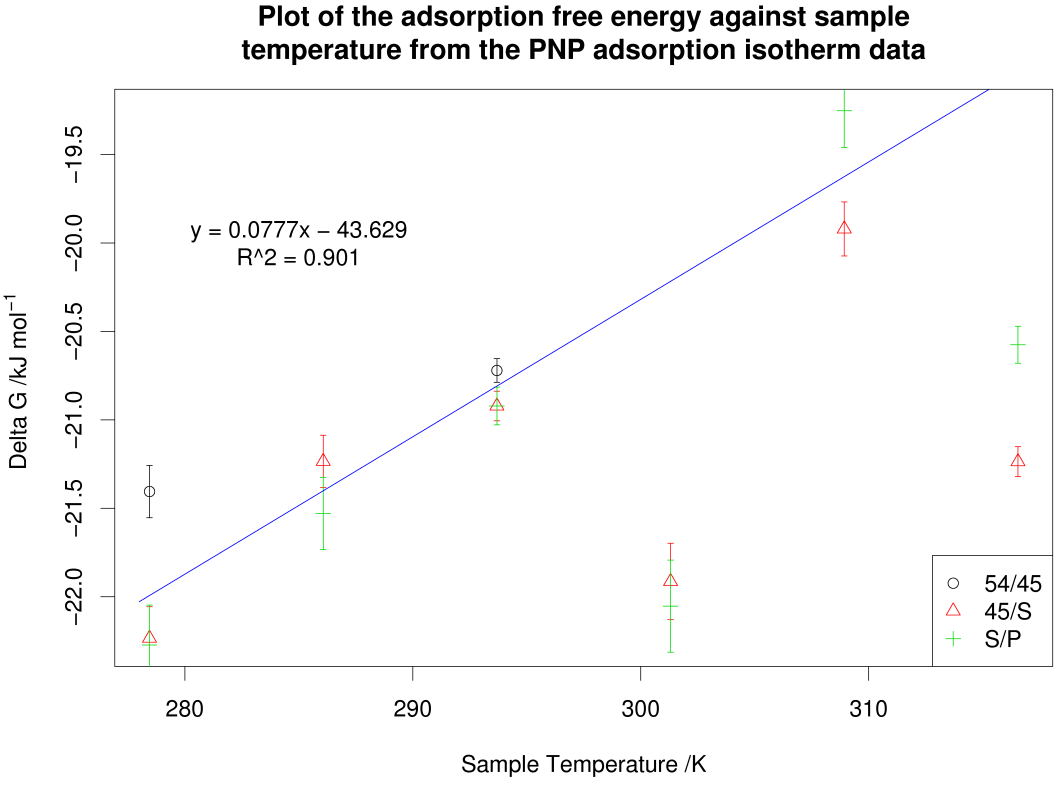


FIGURE 8.21: A plot of surface adsorption free energy from the PNP adsorption isotherm dataset with trend line, assuming the 30 and 50^{o}C data points are outliers

As discussed above, Tamburello-Luca *et al.* re-evaluated a surface tension adsorption isotherm collected by Paluch *et al.*[115]. The details of this process are poorly reported and therefore little trust is given to the reported values. This data abstraction and re-fitting was investigated further to allow comparison of the coverages determined via SHG that calculated through surface tension. The PNP surface tension data was abstracted from the literature[114] using graph digitisation software, collecting the data at 20, 30 and 40^{o}C separately. Initial review of these results show a deviation of the surface tension of pure water at each temperature to that reported in literature[101]; this is shown in table 8.22. The deviation from the expected values suggests the presence of impurities in the water and will affect the accuracy of the results collected during the study.

Temperature /^{o}C	Literature /mNm^{-1}	Paluch /mNm^{-1}
20	72.75±0.05	72.37
30	71.18±0.05	70.50
40	69.56±0.05	69.04

TABLE 8.22: The deviation of experimental determined values of the surface tension of water collect by Paluch *et al.* to those report in literature

The results abstracted from the Paluch literature article were fitted to a Frumkin isotherm via non-linear least squares regression using an R script, as completed by Tamburello-Luca *et al.*. The results from this fitting are compared to those determined by Tamburello-Luca *et al.* in table 8.23 and are plotted in Fig. 8.22.

Data	Temp. /^{o}C	b	k /mol^{-1}dm^3	Δ_{ads}G /kJmol^{-1}
Tamburello-Luca *et al.*	Assumed 20	1.1±0.1	15.12±1.19	-16.4±0.2
This work	20	0.69±0.05	20.73±1.06	-17.17±0.12
	30	0.80±0.08	17.82±1.38	-17.37±0.19
	40	0.66±0.04	18.05±0.66	-17.98±0.09

TABLE 8.23: The determined parameters from the Frumkin fit of the Paluch PNP surface tension data compared to those of Tamburello-Luca *et al.*

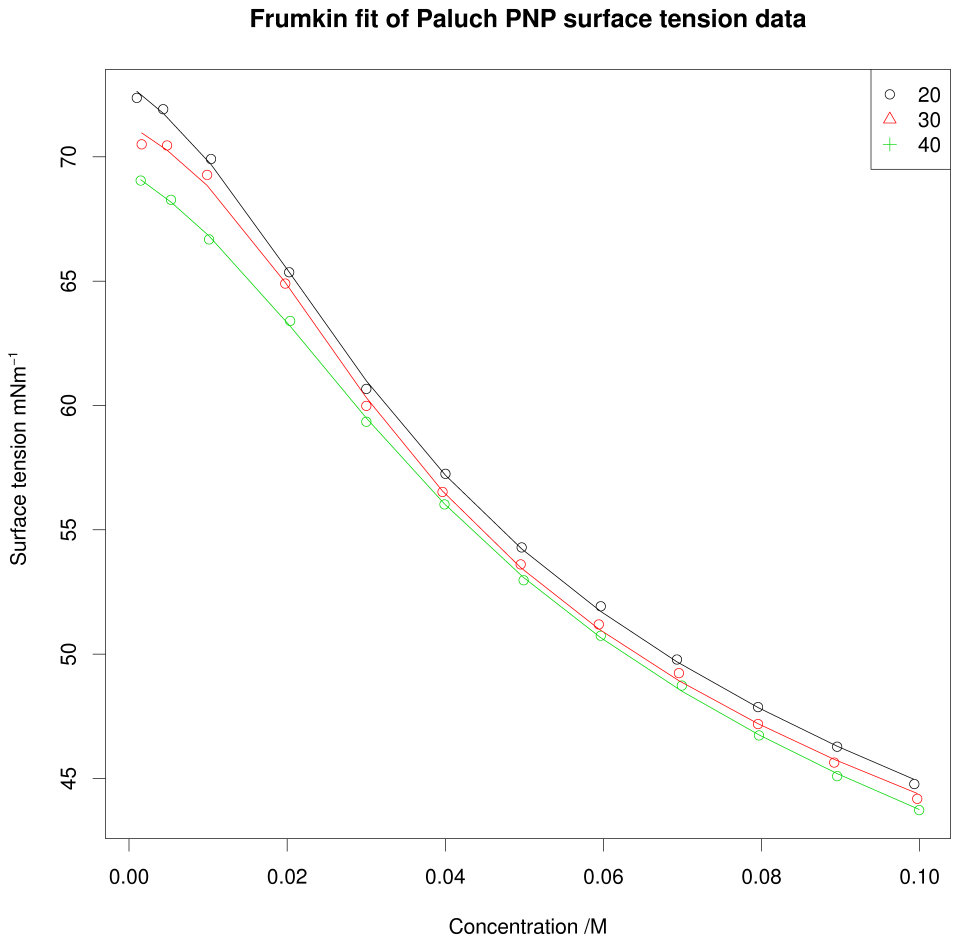


FIGURE 8.22: A plot of Frumkin fit to the Paluch PNP surface tension data

This plot shows a good fit to the surface tension data. The values determined in this work differ significantly to that reported by Tamburello-Luca *et al.*, as the details of the fitting carried out in that paper are not given the discrepancy cannot be investigated further. This PNP surface tension data was further analysed by fitting to a Langmuir isotherm to investigate the significance of the additional Frumkin term in the previous fitting. To complete the fitting the surface tension of pure water, γ_0, was fixed to the literature values before fitting. This was completed via a non-linear least squares regression in R. The results from this are shown in table 8.24 and are plotted in Fig. 8.23.

Temperature /^{o}C	γ_0	k /mol^{-1}dm^3	ΔG /kJmol^{-1}
20	72.75	15.68±3.99	-16.49±0.55
30	71.18	12.64±4.26	-16.51±0.73
40	69.56	11.68±2.81	-16.85±0.56

TABLE 8.24: The determined parameters from the Langmuir fit of the Paluch PNP surface tension data

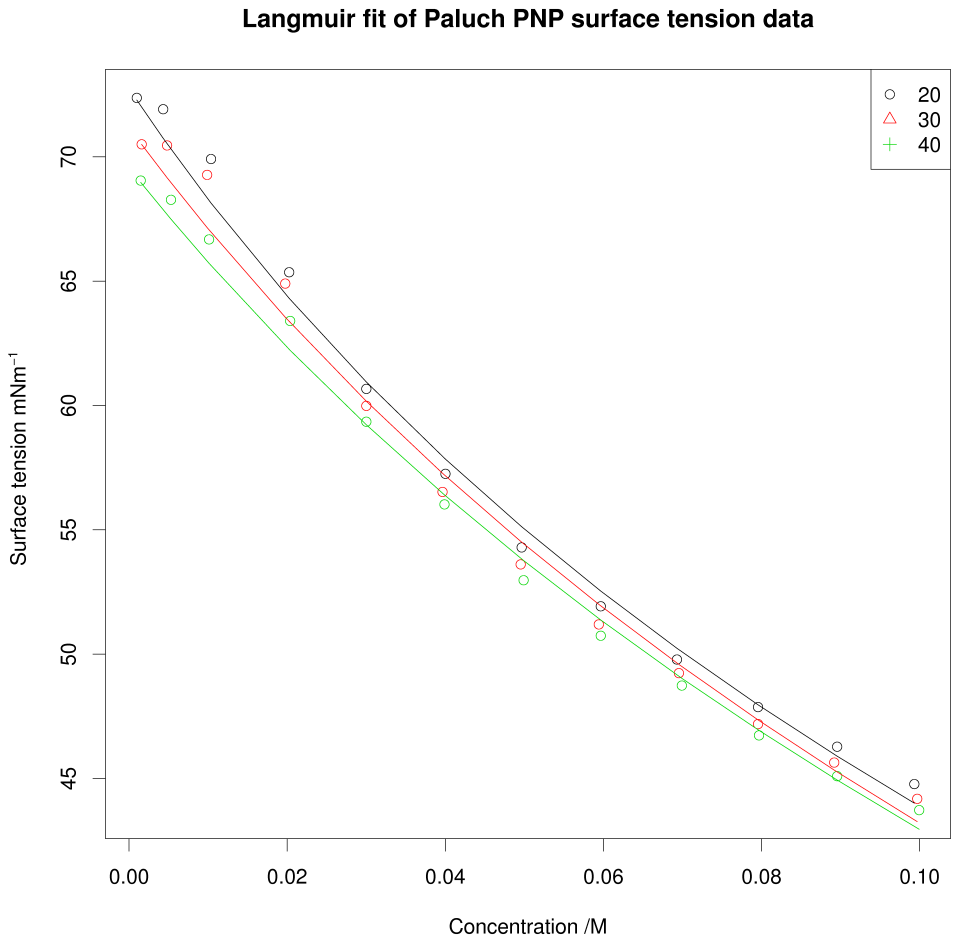


FIGURE 8.23: A plot of Langmuir fit to the Paluch PNP surface tension data

These results do not fit the experimental data as well as the Frumkin isotherm. This may be due to forcing the surface tension of pure water, γ_0, to that of the literature values; it has already been discussed above that the experimental results deviate from these expected values. The results calculated from the Langmuir isotherm show similar values to that of the Tamburello-Luca Frumkin isotherm; this could be investigated further if more information was given in the article.

Using the two isotherms discussed above, the coverages calculated from the surface tension data can be compared to that from the SHG experiments. This comparison is given in table 8.25.

It can be seen that the SHG Langmuir and surface tension Frumkin give comparable coverages over the range of the concentrations. The surface tension Langmuir showing a poor match to those determined via SHG, likely due to the issues in the fitting discussed above. The deviation between SHG and Frumkin coverages may be due to the potential presence of the impurities in the surface tension data, discussed above. The methods

	Coverage /%		
Concentration /mM	SHG	Frumkin	Langmuir
75.12	88	83	54
49.16	82	74	43
29.92	74	57	32
10.80	51	23	14

TABLE 8.25: A comparison of the PNP coverages determined from the SHG adsorption isotherm and the Paluch surface tension data fitted to a Frumkin and Langmuir isotherm

used to collect each dataset differ considerably and therefore introduce different errors in the determined values.

8.4.3 Concentration dependence revisited

Using the information gathered from the adsorption isotherm results, the concentration dependence dataset can be re-worked to determine the molecular orientation parameters omitting the contributions from water that will also be present at the surface. This can be achieved through analysis of the 45 input /S output polarisation combination isotherm which relate to the χ_{xxz} term (therefore take χ=χ_{xxz} in this section unless stated otherwise). The fitting of this data provides the susceptibility of water, χ_{water}, the susceptibility of a monolayer of sample, χ_{sample} and phase between these, ϕ. The susceptibilities can be normalised to water using Eq. 8.6 and 8.7.

$$\chi_{norm,water} = \frac{\chi_{water}}{\chi_{water}} = 1 \tag{8.6}$$

$$\chi_{norm,sample} = \frac{\chi_{sample}}{\chi_{water}} \tag{8.7}$$

The surface susceptibility can then be re-evaluated using Eq. 8.8. The coverage, Θ, can be calculating using the equilibrium constant, k, from the adsorption isotherm fitting.

$$\chi_{norm,surface} = \chi_{norm,water} + \Theta\chi_{norm,sample}e^{i\phi} \tag{8.8}$$

This provided a normalised surface susceptibility, $\chi_{norm,surface}$, at each concentration collected during the adsorption isotherm. At the concentrations corresponding to that collected in the concentration dependence investigation these values can be used to scale the susceptibility terms determined from the polarisation scan (concentration dependence) experiments using Eq. 8.9

$$\chi(Pol.Scan)_{norm,ijk,surface} = \chi(Pol.Scan)_{ijk}\frac{\chi(Ads.iso.)_{norm,xxz,surface}}{\chi(Pol.Scan)_{xxz}} \tag{8.9}$$

where ijk is zxx, xxz and zzz. All three susceptibility terms for water determined in chapter 7 were also normalised to χ_{xxz}=1 to allow them to be used in further analysis.

Using the normalised susceptibility of the surface from the concentration dependence polarisation scans and water, the susceptibility of the sample component omitting water can be calculated using Eq. 8.10

$$\chi_{ijk,sample} = \chi_{norm,ijk,surface} - \chi_{norm,ijk,water} \tag{8.10}$$

where ijk is zxx, xxz and zzz. These sample susceptibility terms can be submitted to Eq. 6.30 and 6.28 as before to determine the orientation parameter, D, and ratio between dominant hyperpolarisability terms of the sample omitting the water contribution. The results from this can be seen in table 8.26 and are plotted in Fig. 8.24.

Conc. /mM	Coverage /%	$\frac{\beta_{zxx}}{\beta_{zzz}}$	D ϕ=90^o	Note
75.12	88	-0.523($\pm$0.007) - 0.128($\pm$0.020)i	0.846($\pm$0.003) - 0.001($\pm$0.007)i	Revised fitting
49.16	82	-0.549($\pm$0.011) - 0.156($\pm$0.027)i	0.853($\pm$0.004) - 0.019($\pm$0.010)i	including water
29.92	74	-0.501($\pm$0.005) + 0.098($\pm$0.019)i	0.855($\pm$0.002) + 0.011($\pm$0.007)i	contribution
10.80	051	-0.438($\pm$0.006) + 0.151($\pm$0.017)i	0.881($\pm$0.003) + 0.004($\pm$0.006)i	
75.12	88	-0.583($\pm$0.012) + 0.112($\pm$0.020)i	0.845($\pm$0.005) + 0.005($\pm$0.006)i	Revised fitting
49.16	82	-0.548($\pm$0.017) - 0.228($\pm$0.015)i	0.841($\pm$0.006) - 0.091($\pm$0.008)i	removing water
29.92	74	-0.512($\pm$0.009) + 0.054($\pm$0.007)i	0.853($\pm$0.003) + 0.020($\pm$0.003)i	contribution
10.80	51	-0.495($\pm$0.008) + 0.007($\pm$0.009)i	0.880($\pm$0.004) + 0.006($\pm$0.003)i	

TABLE 8.26: The molecular orientation parameters determined from the PNP concentration dependence datasets completed at 20^{o}C from both the standard fitting and revised fitting to remove the contributions from water

These results show the contribution from water to the surface orientation parameter, D, is almost zero. The dominant hyperpolarisability terms are modified slightly through this analysis. The terms at the lowest coverage are modified by the largest amount, as expected as the concentration of water at the surface is greater. As in the previous analysis the results show a linear trend with concentration in the surface orientation parameter, D, corresponding to a shift in average molecular tilt angle from 23.18^o at a coverage of 88% to 20.27^o at 51% coverage. These fittings did not take account for any changes in the refractive indices with concentration. It is thought in including these changes the trends in the data will remain unchanged although the determined parameters will be modified, giving a slightly large shift in orientation with coverage.

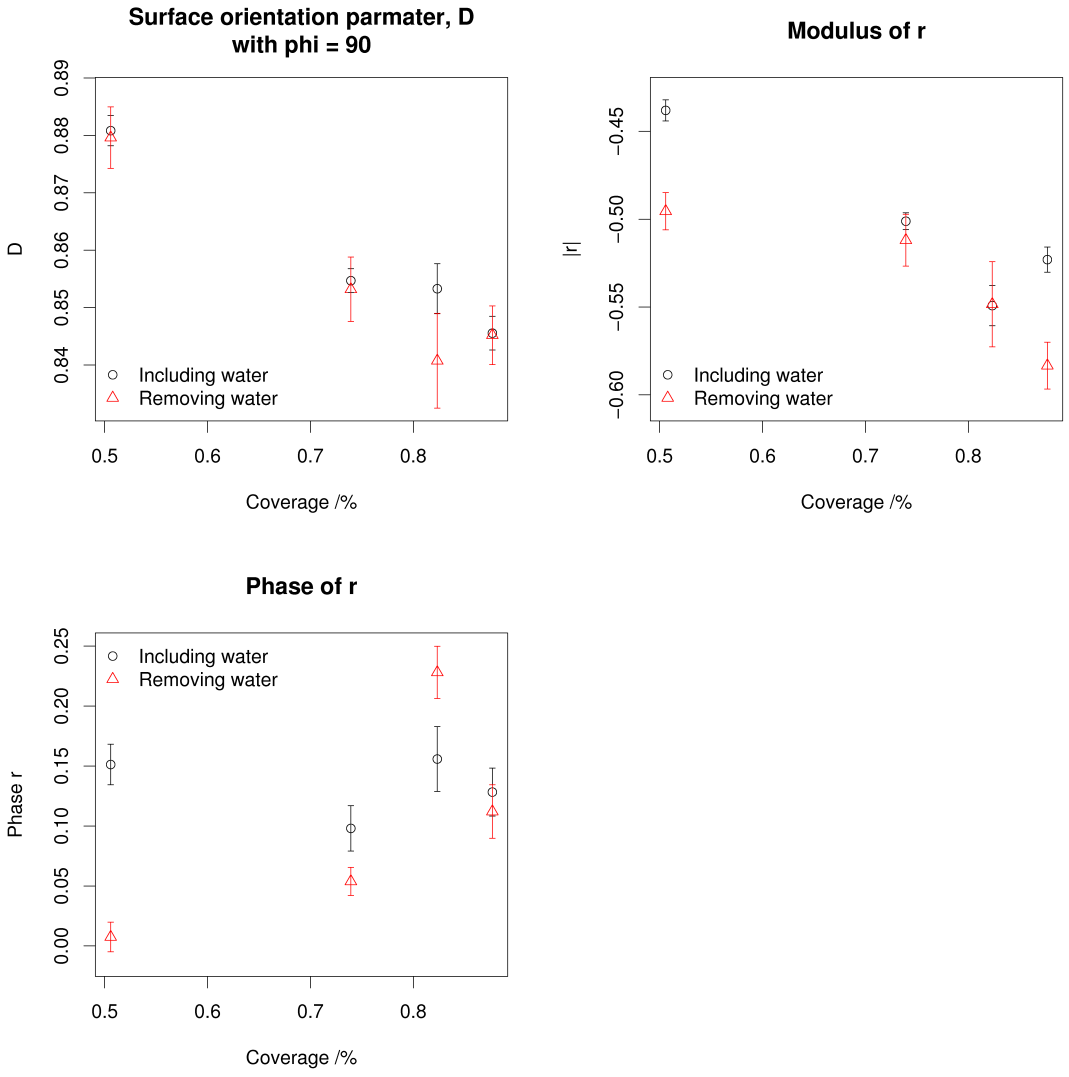


FIGURE 8.24: The plot of the surface orientation parameter, D, and ratio between dominant hyperpolarisability terms against coverage for the PNP concentration dependence datasets run at 20°C from both the standard fitting and revised fitting to remove the contributions from water

8.4.4 Further work

There are several area of this project which can be extended to determine further properties of PNP. Further investigation into the concentration dependence of the surface orientation can be carried out. These can include lower coverages than that studied in this investigation and to look at this dependence at multiple temperatures. To collect data at lower coverage, and therefore lower concentrations, a higher power laser source is required as signal was very low using the existing laser source. The temperature and concentration dependence results have already shown a change in surface orientation as these properties are changed, by collecting data at multiple combinations of these two properties an extended understanding on the interface can be obtained.

Chapter 9

4-Nitrobenzo-15-crown-5

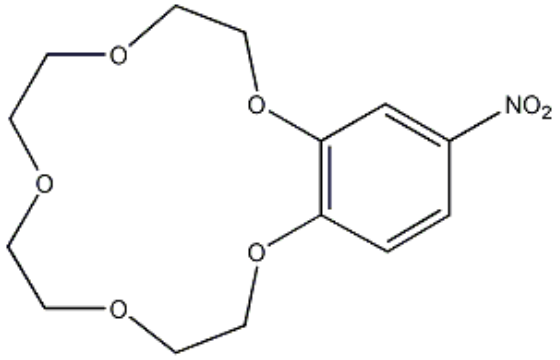


FIGURE 9.1: The molecular structure of NB15C5

Crown ethers are of interest due to their ability to bind strongly to cations. One factor that affects which cation will bond with a given crown is the size of the crown ring. When forming the bound complex there are a number of configurations the system may occupy[123].

1. The cation sits within the ring
2. The cation sits on the ring
3. A 'sandwich' complex is formed with two or more crowns

This project focuses on 15-crown-5 compounds with a crown ring diameter of 1.7 - 2.2 Å[124]. Should a complex described in (1) be produced it would be expected sodium would bind strongly, due to its ionic diameter of 1.9 Å. Previous research using SHG techniques into these crown ethers have shown a change in the observed intensity of the S, 45 and -45 output polarised signal with the addition of sodium, manganese (II) and potassium ions[104] and as such it is believed other complexes may have been produced. Further investigation into this work has suggested there may have been impurities in

the solutions which could have contributed to these results. This project will look to validate and extend this work.

The fitting equations used for water and PNP have been derived for molecules with C2v symmetry. Zhang *et al.*[95] show that for both C2v and Cs molecules there are 3 non-zero independent hyperpolarisability tensor elements with identical interpretations of β_1, β_2 and β_3 used in Eq. 6.27. Therefore Eq. 6.21 can be used to fit the NB15C5 as with the previous data.

Unlike with PNP and water, it is not trivial to define the main molecular axis, z, for NB15C5. Rousay *et al.*[96] discuss *ab initio* calculations to determine the dominant hyperpolarisability terms and have shown the primary axis, O_z, to run from the crown ring to the nitro- group with the O_x perpendicular to this and in the plane of the molecule. This is shown in Fig. 9.2. The SHG results can only be used to determine the orientation of the chromophore section of the molecule and therefore the crown ring is not shown.

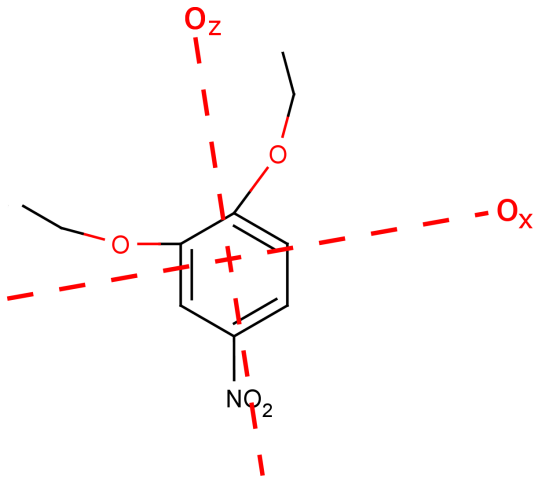


FIGURE 9.2: The molecular axes used for analysing the NB15C5 SHG results

9.1 Experimental

9.1.1 Purification

In previous work[104], off-the-shelf samples of the NB15C5 have been used. These are reported to have a purity of >99.8%. For this work it was decided to purify these samples further in an attempt to provide more repeatable results.

A sample of NB15C5 was dissolved in minimal ethyl acetate and loaded onto a silica column. The eluent was collected until no further compounds were detected via thin layer chromatography (TLC). This process of running the sample through the column

was repeated at least 3 times, leaving a dark yellow stain on the silica. The solvent was removed under vacuum to produce yellow crystals. To aid weighing and transport the crystals were ground to a powder.

After purification, it was possible to produce a solution in water with a maximum concentration of 10 mM. To prepare these samples the solutions were heated to 60^{o}C and sonicated for approximately three hours.

9.1.2 Temperature dependence of surface orientation

As with PNP, a study of the temperature dependence of surface orientation was conducted. The study investigated orientation changes which occur between 0^{o}C and 50^oC. These experiments were conducted using the water bath described in section 6.4.6 and followed the work-flow described in section 6.5.2.1. These experiments were run at water bath temperatures between 0^{o}C and 50^{o}C in 10^{o}C steps. A 7.5 mM NB15C5 solution was used throughout, although a fresh sample was prepared each day.

9.1.3 Effects of salts on surface orientation

To validate the results seen in the previous work a study to determine the effect on the interface when adding a cation was conducted. These experiments were run by addition of the chloride salt of lithium, sodium and potassium during the sample preparation stage. Each salt was run at both a 1x and 10x molar equivalent. A 7.5 mM solution was used throughout and the water bath maintained at 20^{o}C . The workflow described in section 6.5.2.1 was followed to complete the experiments.

9.1.4 Adsorption isotherms

An adsorption isotherm of NB15C5 was completed to determine the surface equilibrium constant, k, at 293K. A 50ml solution of 10 mM NB15C5 underwent 20 sequential dilutions of 5ml pure water. At each concentration the SHG signal was captured over 500 laser shots for polarisation combinations of 45/S and S/P. A second solution of pure water was concentrated by 10 sequential additions of a 10 mM NB15C5 solution. This combination of experiments were chosen to provide adequate overlap of concentrations, allowing comparison between results.

9.2 Results

9.2.1 Temperature dependence on surface orientation

Table 9.1 describes how the dataset containing the results from these experiments was produced. When these experiments were completed there were issues with the reference photodiode and as such no reference signal was captured. Therefore the results discussed in this section will be raw signal rather than normalised. As with the PNP results, there were evaporation issues at the higher concentrations and as such the 50^{o}C data is produced from a number of experiments.

Water bath temp. /^{o}C	Experiment number	Experimental data
0	499	S output
		P output
		45 output
		-45 output
10	498	S output
		P output
		45 output
		-45 output
20	489	S output
		P output
		45 output
		-45 output
30	490	S output
		P output
		45 output
		-45 output
40	491	S output
		P output
		45 output
		-45 output
50	495	S output
	496	P output
	497	45 output
		-45 output

TABLE 9.1: The combination of experiments used to generate the temperature dependence of surface orientation dataset for 7.5 mM NB15C5. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

This dataset was fitted via non-linear regression to Eq. 6.21 using an R script. The model was set up to consider A real with B and C complex. As with the previous

results a $+3^o$ offset was applied to the output polarisation for all data. After an initial fitting it was seen that the 45 output polarisation data varied significantly from the fitted model and as such affected the fit of the other data. The data was re-fitted omitting the 45 polarisation results, giving a much better fit. It is believed the discrepancy of the 45^o output polarisation data is due to evaporation issues. The SHG signal of a 7.5 mM NB15C5 solution was monitored over a number of hours to determine if the surface took time to settle. During this experiment no change in signal was observed. As longer, more detailed, experiments were conducted for the NB15C5 solutions it is likely that evaporation will have effected the later results. 45^o output polarisation is the last to be acquired, and as such the signal will be most affected in this data.

Table 9.2 shows the results from both fitting process, Fig. 9.3 shows a summary of fits including the 45 polarisation data and Fig. 9.4 shows a summary of fits omitting the 45 polarisation data.

Water bath temp. ^{o}C	A	B	C	RSE	Note
0	0.451±0.007	-0.325(±0.022) - 0.200(±0.034)i	0.614(±0.016) + 0.270(±0.035)i	0.014	Including
10	0.452±0.006	-0.378(±0.007) - 0.031(±0.032)i	0.556(±0.020) + 0.300(±0.035)i	0.014	45 polarisation
20	0.467±0.011	-0.341(±0.019) + 0.034(±0.048)i	0.635(±0.036) + 0.270(±0.082)i	0.025	data
30	0.391±0.012	-0.249(±0.024) + 0.050(±0.040)i	0.471(±0.038) + 0.233(±0.073)i	0.022	
40	0.389±0.013	-0.246(±0.022) + 0.044(±0.037)i	0.427(±0.037) + 0.256(±0.058)i	0.021	
50	0.386±0.005	-0.207(±0.017) - 0.129(±0.025)i	0.403(±0.010) + 0.103(±0.038)i	0.009	
0	0.454±0.007	-0.296(±0.024) - 0.235(±0.032)i	0.540(±0.028) + 0.413(±0.034)i	0.013	Omitting
10	0.455±0.007	-0.364(±0.013) - 0.109(±0.046)i	0.505(±0.033) + 0.392(±0.040)i	0.014	45 polarisation
20	0.470±0.009	-0.328(±0.023) - 0.157(±0.052)i	0.534(±0.038) + 0.481(±0.040)i	0.017	data
30	0.395±0.009	-0.254(±0.026) - 0.126(±0.056)i	0.404(±0.041) + 0.382(±0.040)i	0.016	
40	0.400±0.007	-0.222(±0.020) - 0.149(±0.034)i	0.306(±0.029) + 0.434(±0.019)i	0.011	
50	0.390±0.005	-0.235(±0.021) + 0.096(±0.050)i	0.414(±0.005) - 0.003(±0.130)i	0.009	

TABLE 9.2: Non-linear fitting values determined from the 7.50 mM NB15C5 temperature dependence dataset after applying a $+3^{o}$ offset to the output polarisation. Results have been report both including and omitting the 45^{o} polarisation data

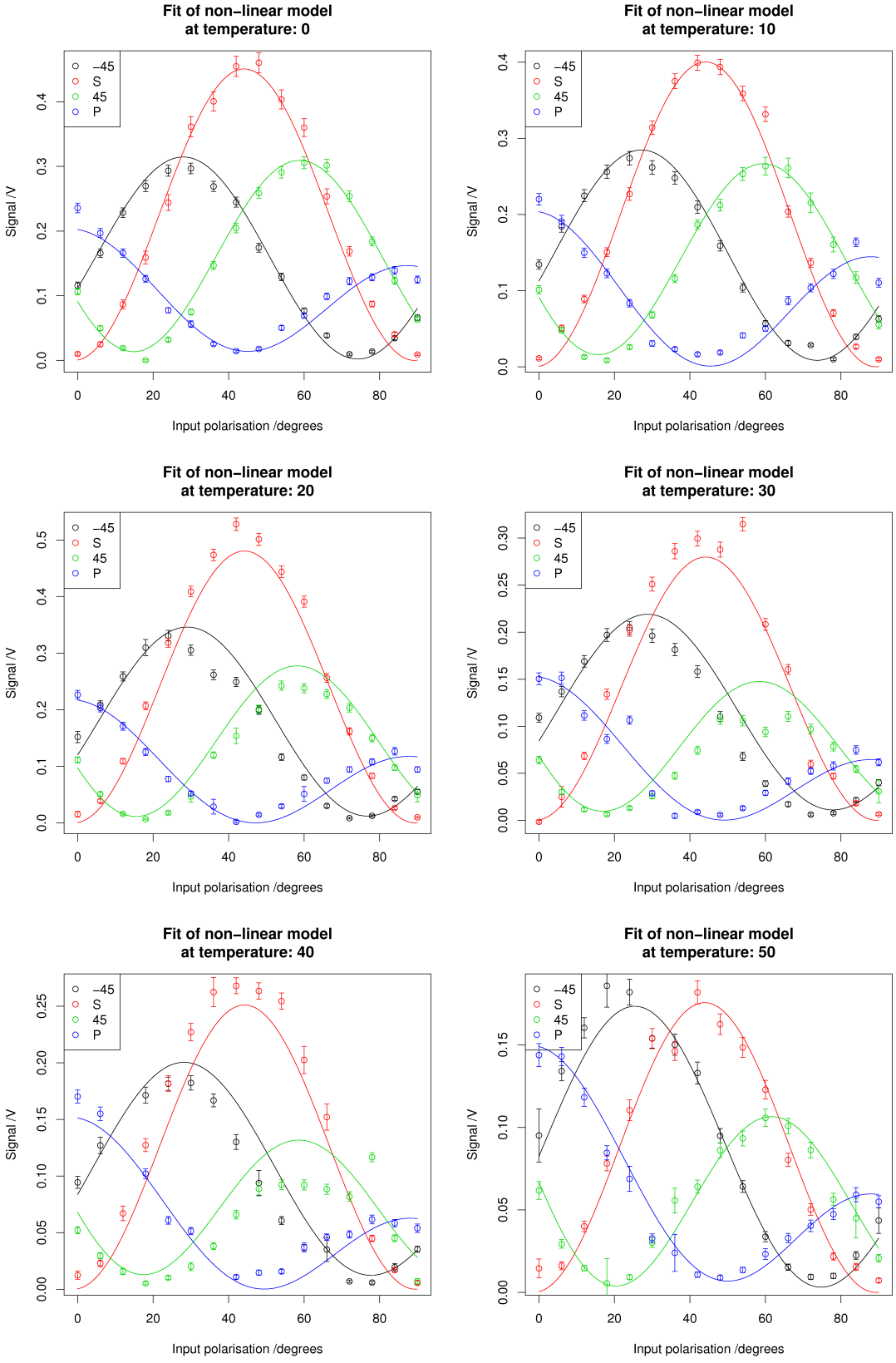


FIGURE 9.3: The summary plot of fitted curves against experimental data for each temperature from 7.50 mM NB15C5 temperature dependence dataset after applying the -3^o offset to the output polarisation. 45 polarisation data is included during the fitting process. Water bath temperatures are given. Error bars are $\pm 3\sigma$

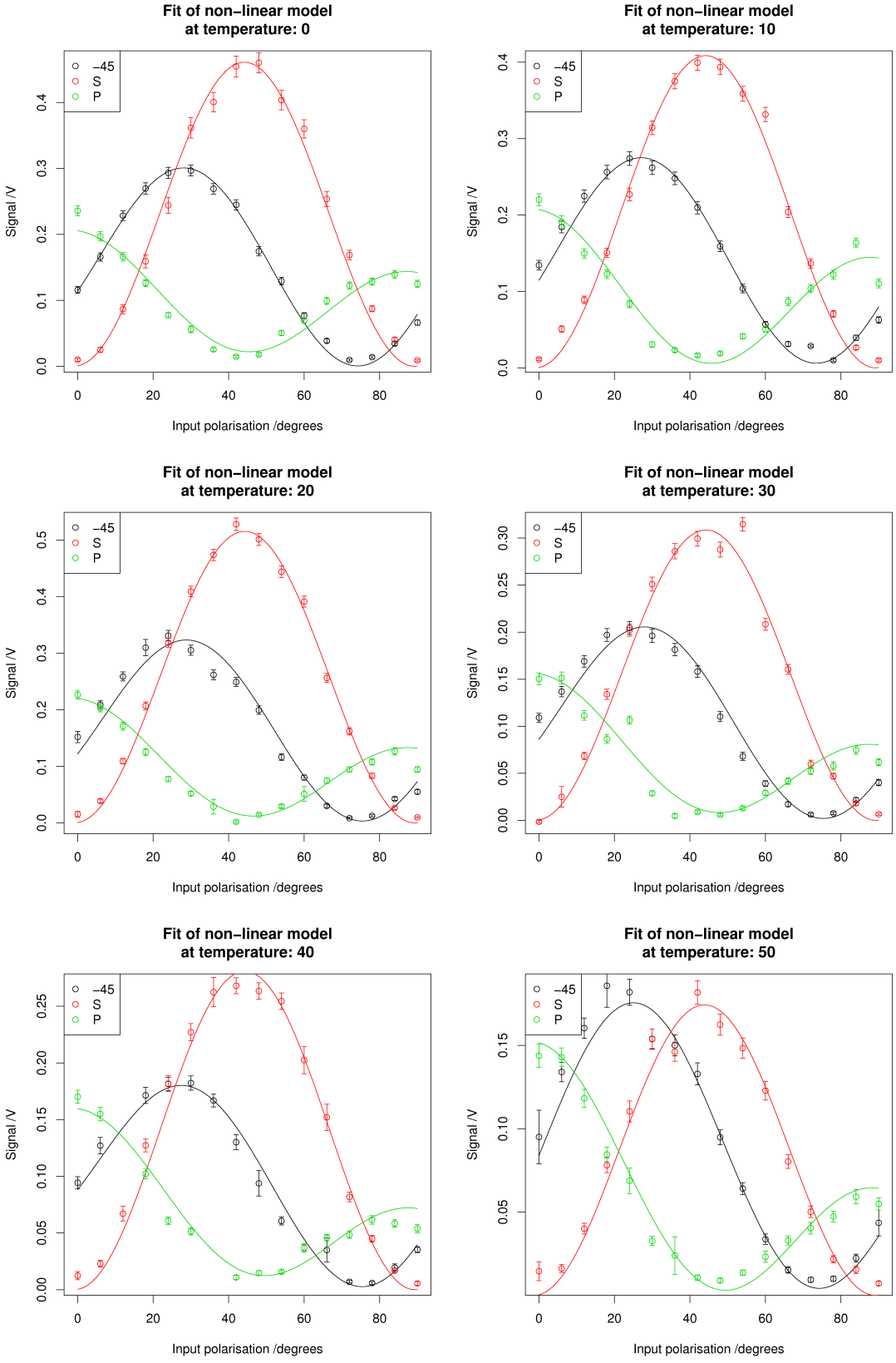


FIGURE 9.4: The summary plot of fitted curves against experimental data for each temperature from 7.50 mM NB15C5 temperature dependence dataset after applying the -3^o offset to the output polarisation. 45 polarisation data is omitted during the fitting process. Water bath temperatures are given. Error bars are $\pm 3\sigma$

9.2.2 Effects of salts on surface orientation

Table 9.3 describes the salt dependence dataset.

Molar equiv of salt	Experiment number
1x LiCl	502
10x LiCl	503
1x NaCl	504
10x NaCl	508
1x KCl	510
10x KCl	511
None	489

TABLE 9.3: The combination of experiments used to generate the salt dependence of surface orientation dataset for 7.5 mM NB15C5. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

As with the previous experiments, the results were fitted to Eq. 6.21 via an R script and the 45^o output polarisation data was omitted. Table 9.4 shows the values of the fitted coefficients and Fig. 9.5 shows a summary plot of fitted values.

Molar equiv of salt	A	B	C	Residual standard error
1x LiCl	0.317±0.006	-0.207(±0.022) - 0.121(±0.032)i	0.331(±0.023) + 0.317(±0.026)i	0.008
10x LiCl	0.324±0.010	-0.211(±0.036) - 0.125(±0.049)i	0.289(±0.041) + 0.345(±0.038)i	0.014
1x NaCl	0.391±0.009	-0.277(±0.033) - 0.147(±0.052)i	0.400(±0.037) + 0.379(±0.043)i	0.015
10x NaCl	0.294±0.005	-0.175(±0.019) - 0.095(±0.030)i	0.277(±0.018) + 0.215(±0.026)i	0.006
1x KCl	0.360±0.007	-0.259(±0.015) - 0.045(±0.049)i	0.404(±0.030) + 0.245(±0.052)i	0.011
10x KCl	0.364±0.007	-0.276(±0.027) - 0.205(±0.033)i	0.380(±0.027) + 0.379(±0.029)i	0.011
None	0.470±0.009	-0.328(±0.029) - 0.157(±0.049)i	0.534(±0.035) + 0.481(±0.042)i	0.017

TABLE 9.4: Non-linear fitting values determined from the 7.5 mM NB15C5 salt dependence dataset after applying a $+3^o$ offset to the output polarisation. The 45^o output polarisation data has been omitted from the fit.

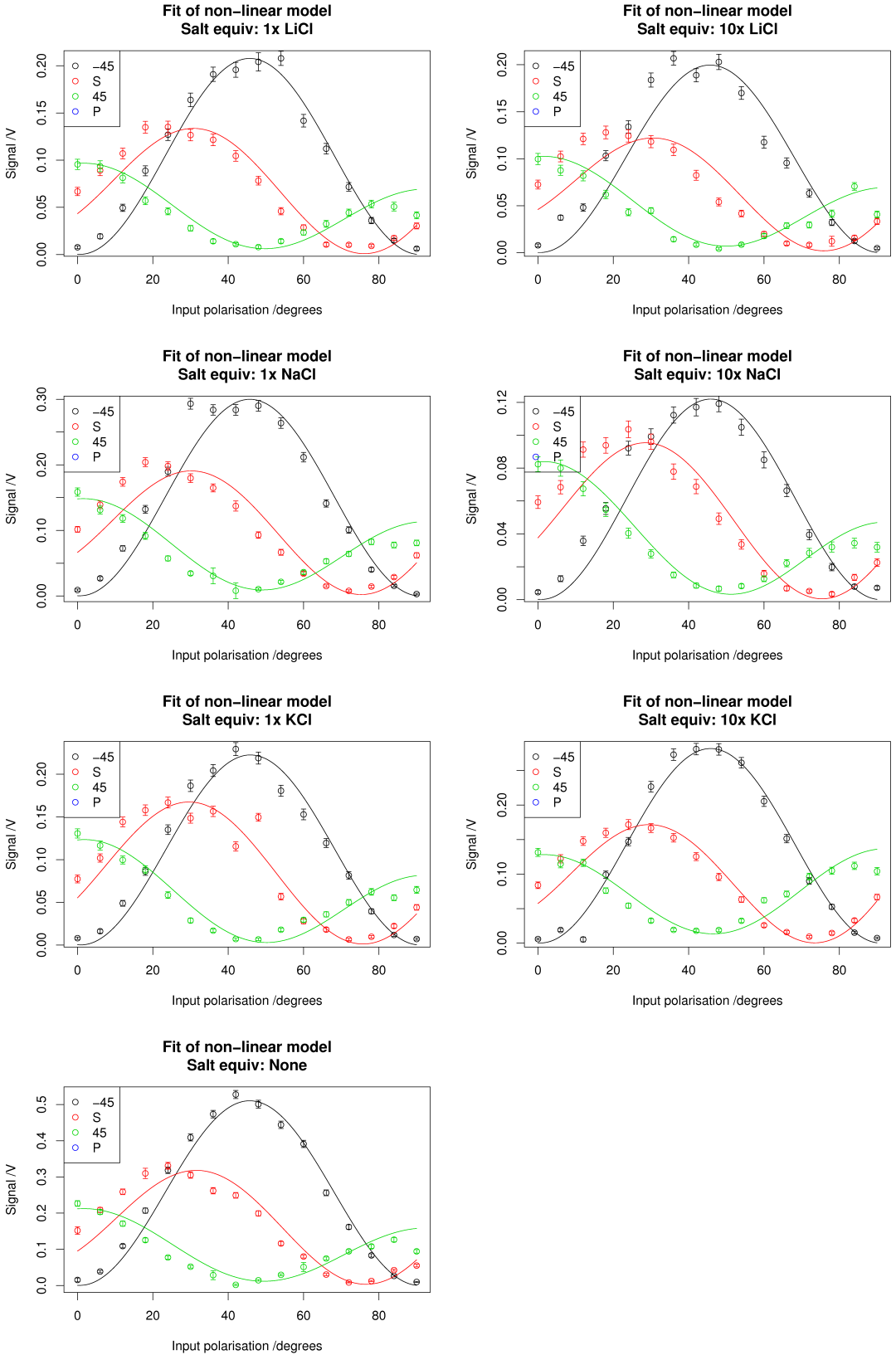


FIGURE 9.5: The summary plot of fitted curves against experimental data for the salt dependence 7.5 mM NB15C5 data after applying the $+3^o$ offset to the output polarisation. Error bars are $\pm 3\sigma$

9.2.3 Adsorption isotherms

As the NB15C5 solution started at a lower concentration to that of the PNP the dilution range could be completed in a single experiment. This means the dataset contains results from only two experiments, as shown in table 9.5.

Isotherm part	Experiment number	Concentration range /mM
Dilution	521	10.21 - 0.30
Concentration	516	0 - 6.17

TABLE 9.5: The combination of experiments used to generate the adsorption isotherm dataset for NB15C5 at 20^{o}C . The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

These results were fitted to Eq. 6.47 through an R script. The script used a non-linear minimisation on the residuals to fit the parameters, allowing $\chi_{solvent}$, χ_{sample} and the equilibrium constant, k, to be fitted. Table 9.6 shows these fitted parameters and Fig. 9.6 shows a plot of the fit.

Polarisations	$\chi_{solvent}$	χ_{sample}	$\cos(\phi)$	k /mol^{-1}dm^3
S/P	0.088±0.003	0.251±0.073	0.999	57.657±26.030
45/S	0.039±0.005	0.351±0.015	0.139	148.196±12.585

TABLE 9.6: Non-linear fitting values determined from the NB15C5 adsorption isotherm dataset at 20^{o}C

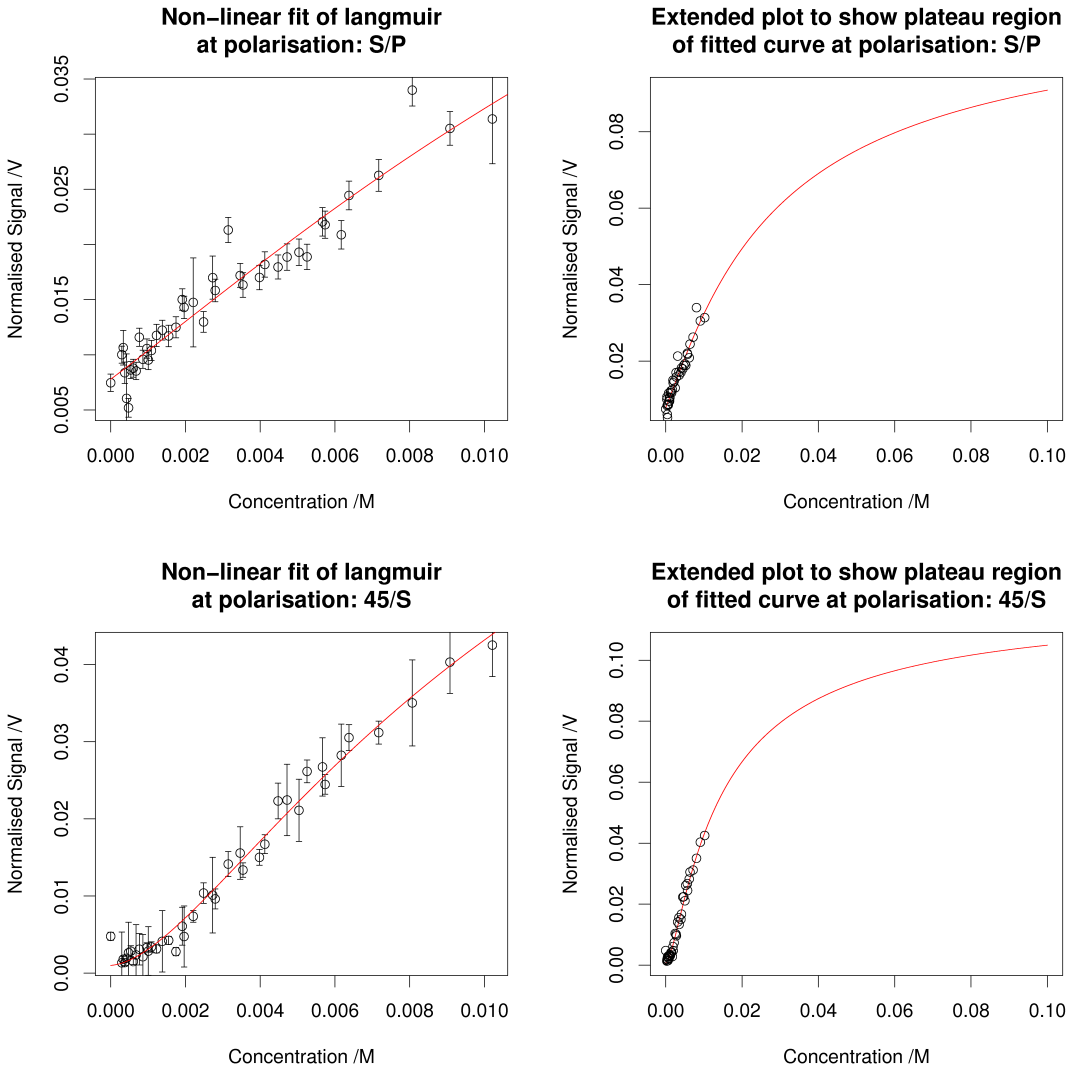


FIGURE 9.6: The summary plot of fitted curves against experimental data from NB15C5 adsorption isotherm dataset at 20^{o}C . Error bars are $\pm 3\sigma$

As shown in Fig. 9.6 the experimental data does not include a plateau region at the higher concentrations, which will affect the quality of the fit. When completing the experiments the maximum concentration available within reasonable time scales was used as the starting concentration. In an attempt to start the experiment at a higher concentration a 20 mM NB15C5 solution was prepared by heating the sample to 70^{o}C and sonicating for 4 hours. As the sample cooled back to room temperature the NB15C5 precipitated out of solution, it was concluded that it was not possible to use a higher concentration at room temperature. As there is no plateau region in the experimental data, the fitted values of k will be poorly determined.

9.3 Discussion

9.3.1 Temperature dependence of surface orientation

To relate the fitted coefficients to surface orientation, the ratios between B/A, C/A and B/C were determined. These ratios related to terms of the susceptibility, which is used in further calculations to determine the molecular orientation. Table 9.7 shows these ratios for the NB15C5 temperature dependence results and Fig 9.7 shows a plot of this data. The temperature reported in these results have been converted from water bath temperature to sample temperature as described in section 6.4.6. The values determined when omitting 45 output polarisation have been used.

Sample temp. /K	B/A	C/A	B/C
278.5	0.832±0.040	1.498±0.021	0.556±0.018
286.1	0.834±0.043	1.404±0.023	0.594±0.022
293.7	0.774±0.050	1.530±0.023	0.506±0.022
301.3	0.718±0.082	1.407±0.039	0.510±0.042
308.9	0.670±0.061	1.328±0.028	0.505±0.035
316.6	0.651±0.058	1.061±0.030	0.614±0.052

TABLE 9.7: The ratio between fitted values from the 7.5 mM NB15C5 temperature dependence datasets

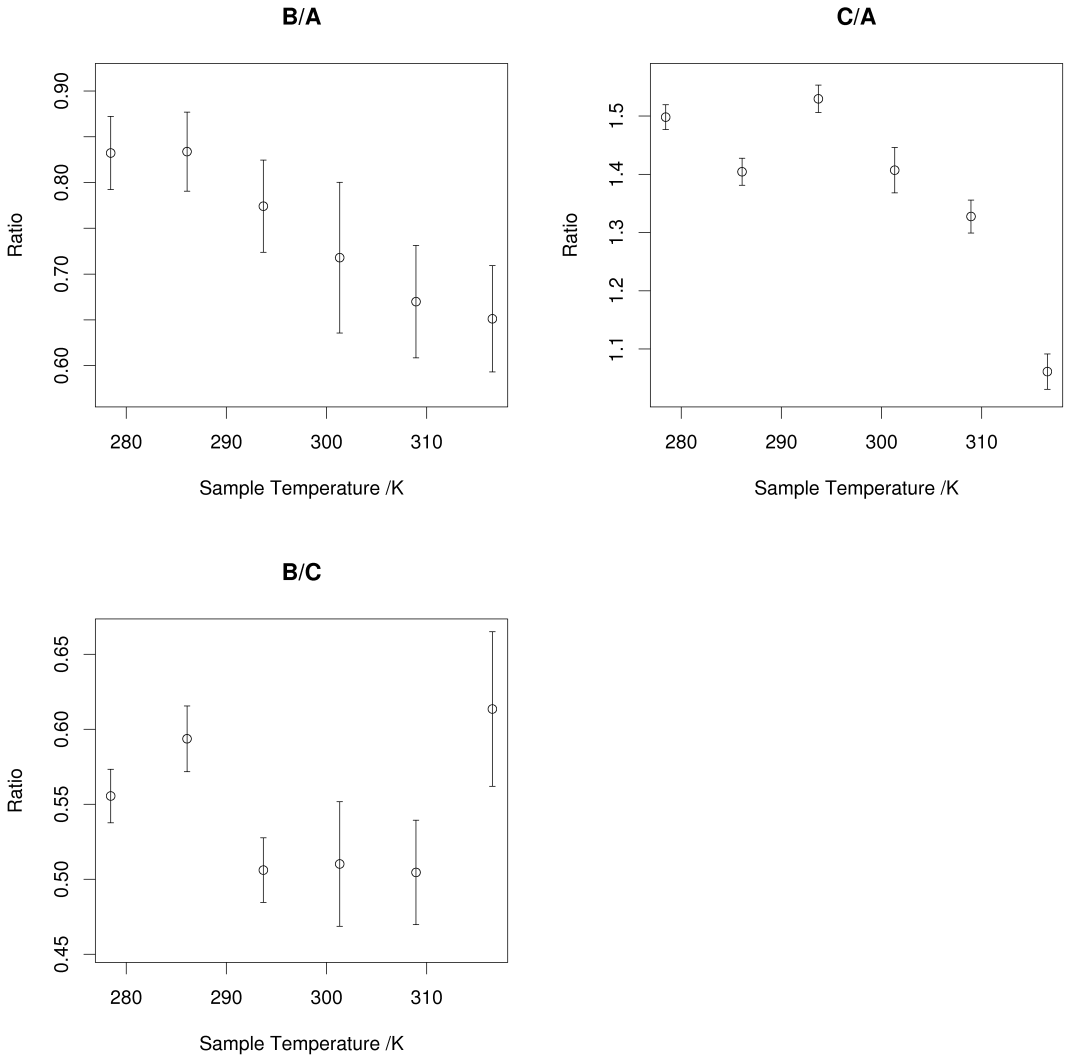


FIGURE 9.7: The plot of ratios between fitted values for the 7.5 mM NB15C5 temperature dependence datasets

These results show a steady change in the ratios between B/A and C/A ratios with temperature, suggesting the orientation changes steadily. This is unlike what was seen with the PNP results, see Fig. 8.9, where a transition occurred at approximately 24^{o}C .

These experimental parameters were submitted to Eq. 6.28 and 6.29/6.30 to determine the molecular tilt angle. To complete this fitting, an estimate of the refractive indices at the interface, n_3, is required. As with the PNP these were determined via spectroscopic ellipsometry. A 7.5 mM NB15C5 was loaded into the ellipsometer and the angle of incidence set to 53^o . A spectroscopic experiment was carried out, scanning between wavelengths of 300 nm and 700 nm in 5 nm steps. The results were fitted to an Quatre Oscillator dispersion model using three oscillators. The first oscillator was seeded with values relating to the dense layer of water as discussed with PNP, section 8.4.1.1. The remaining two oscillators were seeded with oscillator strengths obtained from the UV-VIS spectrum of NB15C5. When fitting these results the model converged to the oscillator strengths seen in the UV-VIS and an angle of incidence set in the experiments. A plot of the fitted curve against experimental values can be seen in Fig. 9.8. The fitting of these results provided an estimate of the refractive index at fundamental and second harmonic wavelength as 2.006 and 2.217 respectively. These differ considerably from those used in previous studies (the average between those determined for air and water).

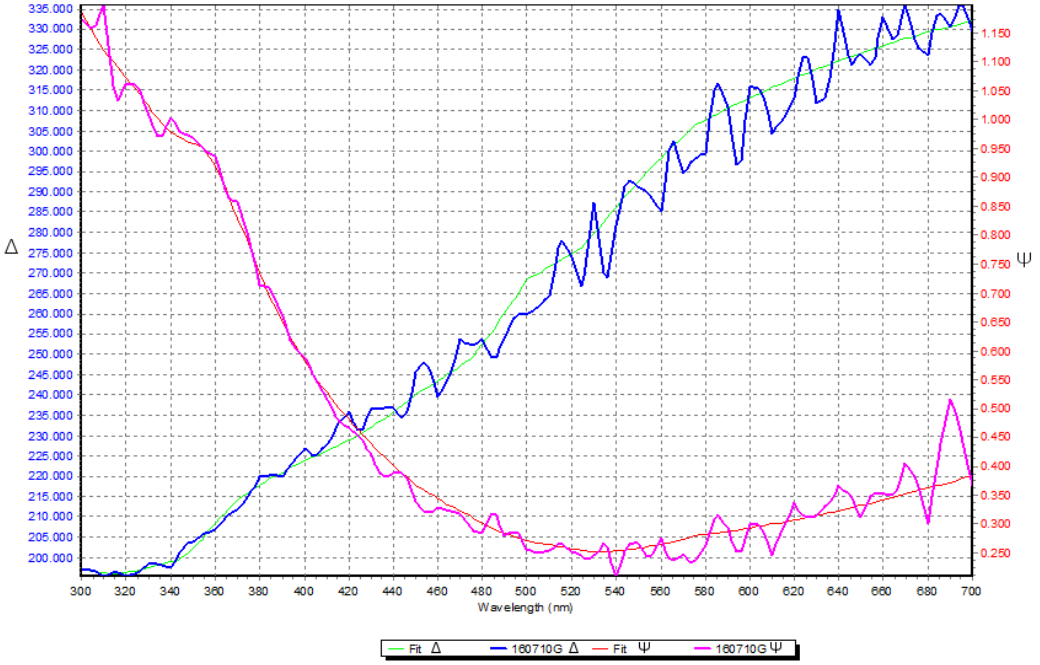


FIGURE 9.8: The plot of fitted curves against experimental results for the 7.5 mM NB15C5 spectroscopic ellipsometry experiment

The experimental results and refractive indices were used to calculate the ratio between dominant terms in the hyperpolarisability and the surface orientation parameter, D. These results are shown in table 9.8 and 9.9 with the parameters determined using the

average refractive index values between the two interfaces (1.165 at both wavelengths) have been included for comparison. A plot of these results can be seen in Fig. 9.9, for ellipsometry determined refractive indices, and Fig. 9.10, for average refractive indices.

Refractive index	Sample temp. /K	χ_{xxz}	χ_{zxx}	χ_{zzz}
Ellipsometry Values	278.45	6.122(±0.181) + 4.683(±0.401)i	-5.448(±0.404) - 4.327(±0.630)i	221.233(±4.360) + 100.889(±8.345)i
	286.07	5.727(±0.225) + 4.445(±0.401)i	-6.699(±0.124) - 2.008(±0.582)i	224.381(±4.248) + 81.015(±7.862)i
	293.69	6.053(±0.411) + 5.449(±0.930)i	-6.041(±0.342) - 2.893(±0.886)i	227.534(±7.665) + 102.337(±17.444)i
	301.31	4.576(±0.428) + 4.328(±0.825)i	-4.674(±0.433) - 2.328(±0.744)i	180.718(±7.941) + 81.506(±15.290)i
	308.93	3.471(±0.417) + 4.918(±0.662)i	-4.096(±0.407) - 2.753(±0.678)i	161.044(±7.865) + 93.367(±12.389)i
	316.55	4.689(±0.118) - 0.030(±0.428)i	-4.329(±0.322) + 1.765(±0.465)i	178.982(±3.186) - 12.857(±8.572)i
Average Values	278.45	2.348(±0.061) + 1.033(±0.135)i	-1.654(±0.112) - 1.017(±0.174)i	7.445(±0.137) + 2.261(±0.262)i
	286.07	2.127(±0.076) + 1.146(±0.135)i	-1.924(±0.034) - 0.157(±0.161)i	7.358(±0.133) + 1.733(±0.247)i
	293.69	2.429(±0.139) + 1.034(±0.314)i	-1.731(±0.094) + 0.174(±0.245)i	7.720(±0.241) + 1.310(±0.548)i
	301.31	1.802(±0.144) + 0.889(±0.278)i	-1.265(±0.120) + 0.255(±0.205)i	5.995(±0.249) + 1.043(±0.480)i
	308.93	1.633(±0.141) + 0.978(±0.223)i	-1.252(±0.112) + 0.224(±0.187)i	5.737(±0.247) + 1.191(±0.389)i
	316.55	1.540(±0.040) + 0.392(±0.144)i	-1.052(±0.089) - 0.656(±0.128)i	5.427(±0.100) + 1.074(±0.269)i

TABLE 9.8: The molecular susceptibility determined from the 7.5 mM NB15C5 temperature dependence datasets, calculated using both ellipsometry determined and average refractive indices

Refractive index	Sample temp. /K	$\frac{\beta_{zxx}}{\beta_{zzz}}$	D ψ=90^o	D ψ random
Ellipsometry	278.45	-0.111($\pm$0.001) - 0.025($\pm$0.004)i	0.945($\pm$0.001) - 0.011($\pm$0.002)i	0.947($\pm$0.001) - 0.010($\pm$0.002)i
Values	286.07	-0.110($\pm$0.002) - 0.013($\pm$0.003)i	0.948($\pm$0.001) - 0.016($\pm$0.002)i	0.950($\pm$0.001) - 0.015($\pm$0.002)i
	293.69	-0.109($\pm$0.004) - 0.018($\pm$0.003)i	0.944($\pm$0.002) - 0.016($\pm$0.003)i	0.947($\pm$0.001) - 0.015($\pm$0.003)i
	301.31	-0.107($\pm$0.005) - 0.019($\pm$0.003)i	0.946($\pm$0.002) - 0.017($\pm$0.004)i	0.948($\pm$0.002) - 0.016($\pm$0.004)i
	308.93	-0.106($\pm$0.006) - 0.026($\pm$0.003)i	0.946($\pm$0.002) - 0.022($\pm$0.004)i	0.949($\pm$0.002) - 0.021($\pm$0.003)i
	316.55	-0.097($\pm$0.003) + 0.012($\pm$0.003)i	0.953($\pm$0.001) - 0.003($\pm$0.002)i	0.955($\pm$0.001) - 0.003($\pm$0.002)i
Average	278.45	-0.692($\pm$0.008) - 0.091($\pm$0.022)i	0.705($\pm$0.002) - 0.011($\pm$0.007)i	0.765($\pm$0.002) - 0.004($\pm$0.007)i
Values	286.07	-0.692($\pm$0.008) + 0.015($\pm$0.017)i	0.712($\pm$0.002) - 0.048($\pm$0.008)i	0.771($\pm$0.002) - 0.039($\pm$0.007)i
	293.69	-0.651($\pm$0.016) + 0.038($\pm$0.023)i	0.698($\pm$0.004) - 0.047($\pm$0.013)i	0.758($\pm$0.004) - 0.040($\pm$0.011)i
	301.31	-0.624($\pm$0.024) + 0.051($\pm$0.025)i	0.700($\pm$0.007) - 0.058($\pm$0.016)i	0.758($\pm$0.006) - 0.050($\pm$0.013)i
	308.93	-0.623($\pm$0.024) + 0.050($\pm$0.023)i	0.703($\pm$0.007) - 0.067($\pm$0.015)i	0.761($\pm$0.006) - 0.057($\pm$0.013)i
	316.55	-0.633($\pm$0.012) - 0.108($\pm$0.024)i	0.723($\pm$0.003) + 0.002($\pm$0.009)i	0.777($\pm$0.003) + 0.007($\pm$0.008)i

TABLE 9.9: The molecular orientation parameters determined from the 7.5 mM NB15C5 temperature dependence datasets, calculated using both ellipsometry determined and average refractive indices

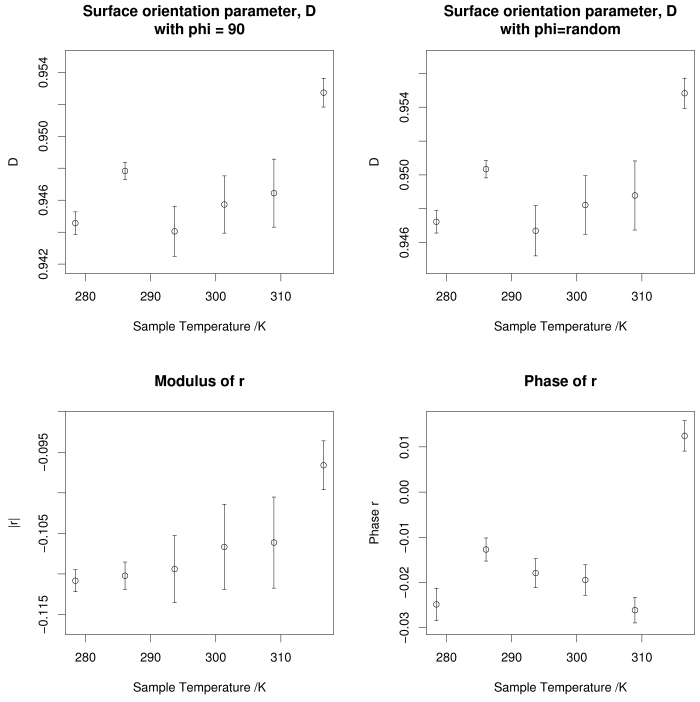


FIGURE 9.9: The plot of surface orientation parameter, D, and ratio between dominant hyperpolarisability terms against sample temperature for the 7.5 mM NB15C5 temperature dependence datasets using refractive indices determined through ellipsometry

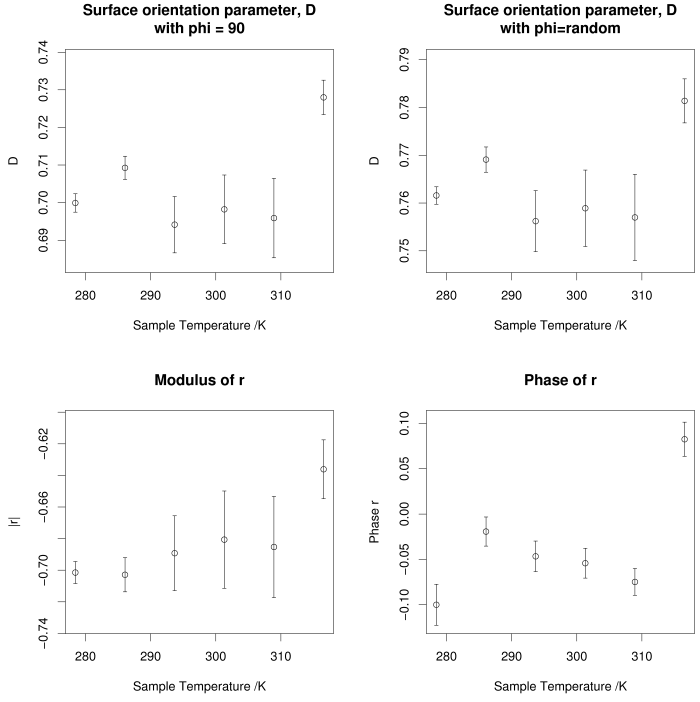


FIGURE 9.10: The plot of surface orientation parameter, D, and ratio between dominant hyperpolarisability terms against sample temperature for the 7.5 mM NB15C5 temperature dependence datasets using the average between the refractive indices of the two media

From these two plots it can be seen that although the determined parameters are significantly changed when modifying the refractive indices of the interface, the trends in the data are not. This shows the importance of accurately determining the refractive indices when reporting absolute values, rather than relative changes. When using the refractive indices determined via ellipsometry the average molecular tilt angle is determined as 13.6^o which remains constant, within experimental error, with changes in temperature.

When fitting the experimental data, the results at 45^o output polarisation were omitted from the fitting as it was thought these results were affected by evaporation. Table 9.10 shows the susceptibility terms and table 9.11 shows the ratio between dominant terms in the hyperpolarisability and the surface orientation parameter, D, for both results including and omitting these experimental results. A plot of these results can be seen in Fig. 9.11, omitting 45^o output polarisation data, and Fig. 9.12, including 45^o output polarisation data.

Dataset	Sample temp. /K	χ_{xxz}	χ_{zxx}	χ_{zzz}
Omitting	278.45	6.122(±0.181) + 4.683(±0.401)i	-5.448(±0.404) - 4.327(±0.630)i	221.233(±4.360) + 100.889(±8.345)i
45^o output	286.07	5.727(±0.225) + 4.445(±0.401)i	-6.699(±0.124) - 2.008(±0.582)i	224.381(±4.248) + 81.015(±7.862)i
polarisation	293.69	6.053(±0.411) + 5.449(±0.930)i	-6.041(±0.342) - 2.893(±0.886)i	227.534(±7.665) + 102.337(±17.444)i
data	301.31	4.576(±0.428) + 4.328(±0.825)i	-4.674(±0.433) - 2.328(±0.744)i	180.718(±7.941) + 81.506(±15.290)i
	308.93	3.471(±0.417) + 4.918(±0.662)i	-4.096(±0.407) - 2.753(±0.678)i	161.044(±7.865) + 93.367(±12.389)i
	316.55	4.689(±0.118) - 0.030(±0.428)i	-4.329(±0.322) + 1.765(±0.465)i	178.982(±3.186) - 12.857(±8.572)i
Including	278.45	6.961(±0.181) + 3.062(±0.401)i	-5.991(±0.404) - 3.682(±0.630)i	237.001(±4.360) + 71.970(±8.345)i
45^o output	286.07	6.305(±0.225) + 3.399(±0.401)i	-6.968(±0.124) - 0.570(±0.582)i	234.244(±4.248) + 55.159(±7.862)i
polarisation	293.69	7.201(±0.411) + 3.065(±0.930)i	-6.270(±0.342) + 0.630(±0.886)i	245.762(±7.665) + 41.693(±17.444)i
data	301.31	5.342(±0.428) + 2.637(±0.825)i	-4.583(±0.433) + 0.924(±0.744)i	190.841(±7.941) + 33.191(±15.290)i
	308.93	4.841(±0.417) + 2.899(±0.662)i	-4.536(±0.407) + 0.811(±0.678)i	182.630(±7.865) + 37.919(±12.389)i
	316.55	4.564(±0.118) + 1.162(±0.428)i	-3.811(±0.322) - 2.376(±0.465)i	172.775(±3.186) + 34.187(±8.572)i

TABLE 9.10: The molecular susceptibility determined from the 7.5 mM NB15C5 temperature dependence datasets for both the full results and omitting the 45^o output polarisation data

Dataset	Sample temp. /K	$\frac{\beta_{zxx}}{\beta_{zzz}}$	D ψ=90^o	D ψ random
Omitting	278.45	-0.111($\pm$0.001) - 0.025($\pm$0.004)i	0.945($\pm$0.001) - 0.011($\pm$0.002)i	0.947($\pm$0.001) - 0.010($\pm$0.002)i
45^o output	286.07	-0.110($\pm$0.002) - 0.013($\pm$0.003)i	0.948($\pm$0.001) - 0.016($\pm$0.002)i	0.950($\pm$0.001) - 0.015($\pm$0.002)i
polarisation	293.69	-0.109($\pm$0.004) - 0.018($\pm$0.003)i	0.944($\pm$0.002) - 0.016($\pm$0.003)i	0.947($\pm$0.001) - 0.015($\pm$0.003)i
data	301.31	-0.107($\pm$0.005) - 0.019($\pm$0.003)i	0.946($\pm$0.002) - 0.017($\pm$0.004)i	0.948($\pm$0.002) - 0.016($\pm$0.004)i
	308.93	-0.106($\pm$0.006) - 0.026($\pm$0.003)i	0.946($\pm$0.002) - 0.022($\pm$0.004)i	0.949($\pm$0.002) - 0.021($\pm$0.003)i
	316.55	-0.097($\pm$0.003) + 0.012($\pm$0.003)i	0.953($\pm$0.001) - 0.003($\pm$0.002)i	0.955($\pm$0.001) - 0.003($\pm$0.002)i
Including	278.45	-0.109($\pm$0.001) - 0.020($\pm$0.004)i	0.945($\pm$0.001) - 0.006($\pm$0.002)i	0.948($\pm$0.001) - 0.005($\pm$0.002)i
45^o output	286.07	-0.109($\pm$0.001) - 0.005($\pm$0.003)i	0.948($\pm$0.000) - 0.013($\pm$0.002)i	0.951($\pm$0.000) - 0.012($\pm$0.002)i
polarisation	293.69	-0.104($\pm$0.002) + 0.000($\pm$0.004)i	0.945($\pm$0.001) - 0.012($\pm$0.003)i	0.948($\pm$0.001) - 0.012($\pm$0.003)i
data	301.31	-0.098($\pm$0.004) + 0.002($\pm$0.005)i	0.947($\pm$0.001) - 0.015($\pm$0.004)i	0.949($\pm$0.001) - 0.014($\pm$0.004)i
	308.93	-0.097($\pm$0.004) + 0.000($\pm$0.005)i	0.948($\pm$0.001) - 0.017($\pm$0.004)i	0.950($\pm$0.001) - 0.016($\pm$0.004)i
	316.55	-0.096($\pm$0.002) - 0.020($\pm$0.004)i	0.952($\pm$0.001) - 0.002($\pm$0.002)i	0.954($\pm$0.001) - 0.002($\pm$0.002)i

TABLE 9.11: The molecular orientation parameters determined from the 7.5 mM NB15C5 temperature dependence datasets for both the full results and omitting the 45^o output polarisation data

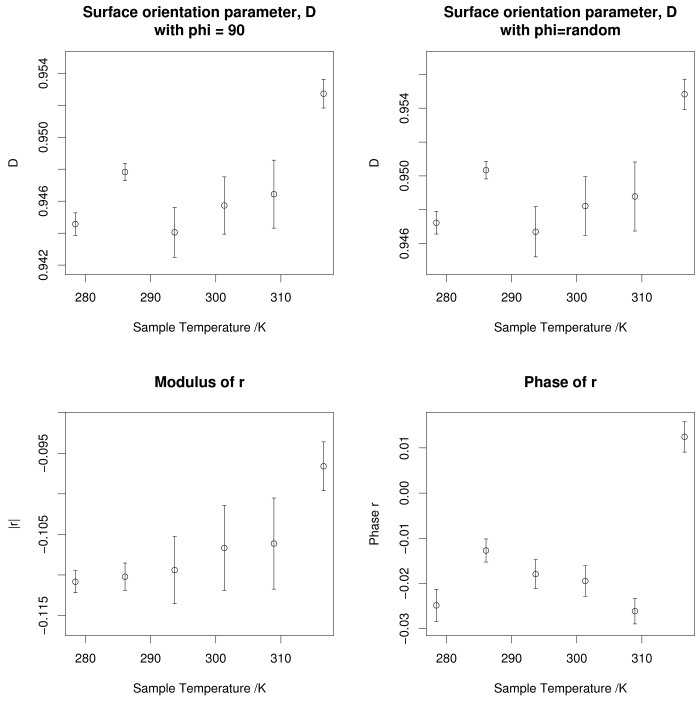


FIGURE 9.11: The plot of surface orientation parameter, D, and ratio between dominant hyperpolarisability terms against sample temperature for the 7.5 mM NB15C5 temperature dependence datasets omitting the 45^o output polarisation data

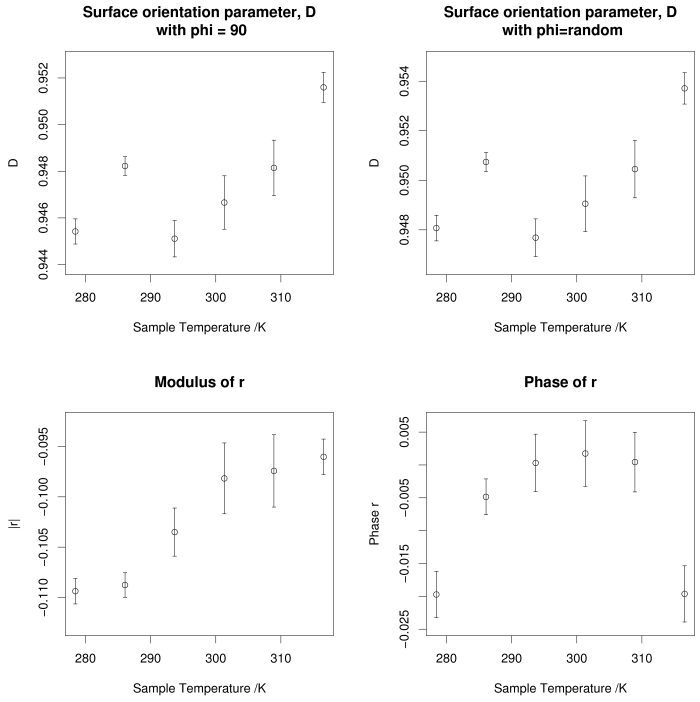


FIGURE 9.12: The plot of surface orientation parameter, D, and ratio between dominant hyperpolarisability terms against sample temperature for the 7.5 mM NB15C5 temperature dependence datasets including the 45^o output polarisation data

Both plots show that there is little change in the surface orientation parameter, D, and therefore molecular tilt angle, θ with changes in temperature. The plot of the $|r|$ when omitting the 45^o polarisation data, Fig. 9.11, may be interpreted as either constant within error or showing a steady change. This change is emphasised in the plot when including the 45^o output polarisation, Fig. 9.12. To determine the correct interpretation of these results, further analysis was carried out.

The experimental data was fitted to the D, F and R equation, Eq. 6.40. It was not possible to fit all parameters directly to the equation as the non-linear regression would fail, therefore the two possible interpretations were investigated. The first using the average values of D and the real and imaginary components of R determined from the experimental data omitting the 45^o output polarisation. The second using the average value determined for D and the imaginary components of R when including the 45 output polarisation experimental data and modelling the change in the real component of R as Re(R) = 0.394E-4*(*sample temp /K*) - 0.220. The experimental results were fitted via an R script using a random distribution of ψ, allowing F to be determined. The results from this fitting is shown in table 9.12. Fig. 9.13 shows a summary plot of the fitting when all values were constant and Fig. 9.14 shows when *Re(R)* was varied.

Model	Sample temp. ^{o} C	*D*	*Re(R)*	*Im(R)*	*F*	RSE
Fixed values	278.5	0.95	-0107	0.019	136.46±1.05	0.024
	286.1				129.50±0.83	0.018
	293.7				138.44±1.63	0.038
	301.3				107.04±1.63	0.029
	308.9				102.26±1.64	0.028
	316.6				91.38±1.64	0.019
Fixed *D*, *Im(R)*	278.5	0.95	-0.110	0.019	136.54±1.00	0.023
Chaning *Re(R)*	286.1		-0.107		129.51±0.82	0.018
	293.7		-0.104		138.26±1.57	0.037
	301.3		-0.101		106.85±1.41	0.026
	308.9		-0.098		102.06±1.35	0.023
	316.6		-0.095		90.56±0.70	0.011

TABLE 9.12: The fitting results from the 7.5 mM NB15C5 temperature dependence dataset fitted to the *F*,*D* and *R* equation, comparing the two interpretations of the *D* and *R* fitting results

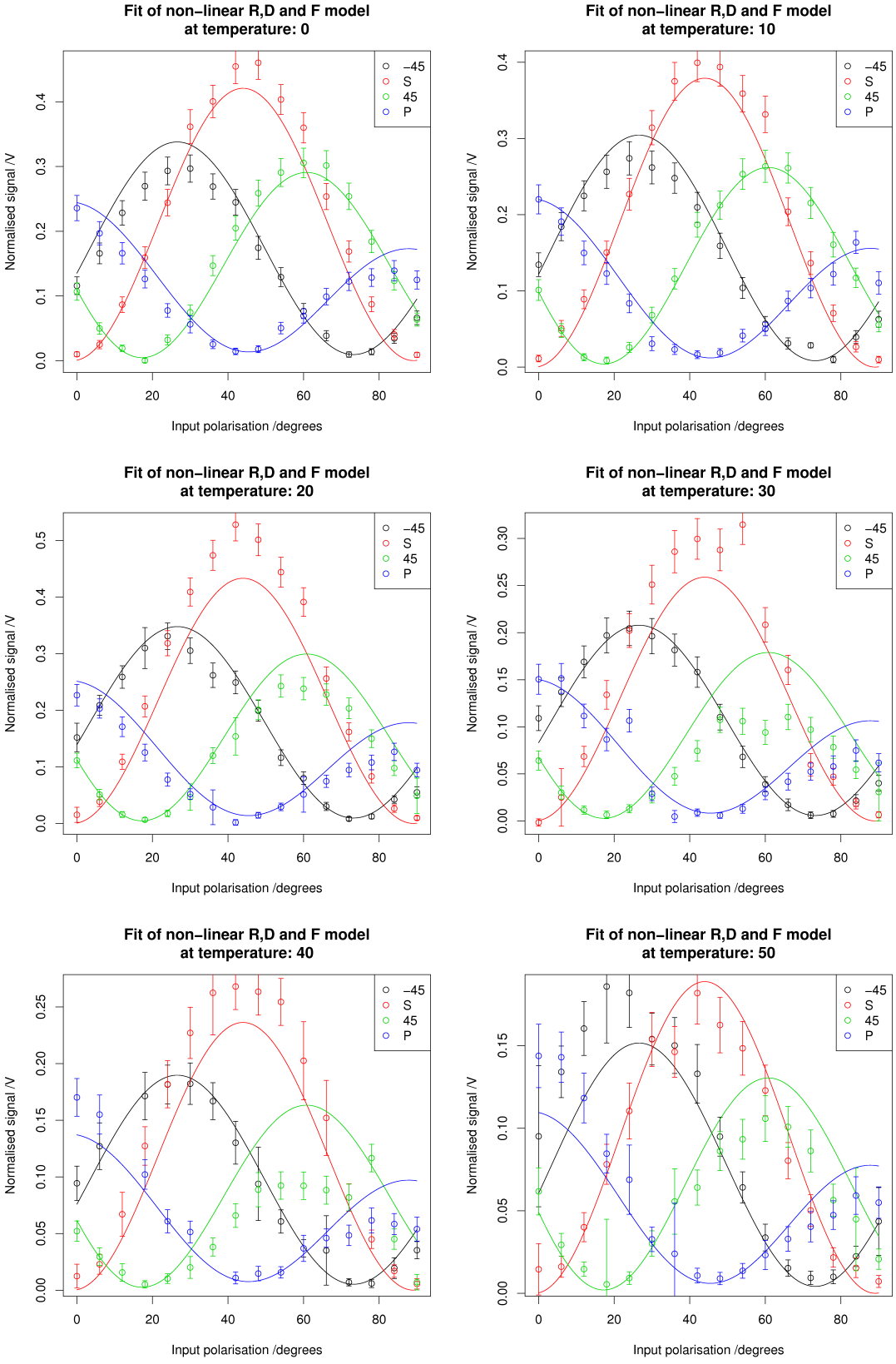


FIGURE 9.13: The summary plot of fitted curves against experimental data for each temperature from the 7.5 mM NB15C5 dataset fitted to the F, D and R model with fixed D, $Re(R)$ and $Im(R)$. Water bath temperatures are given

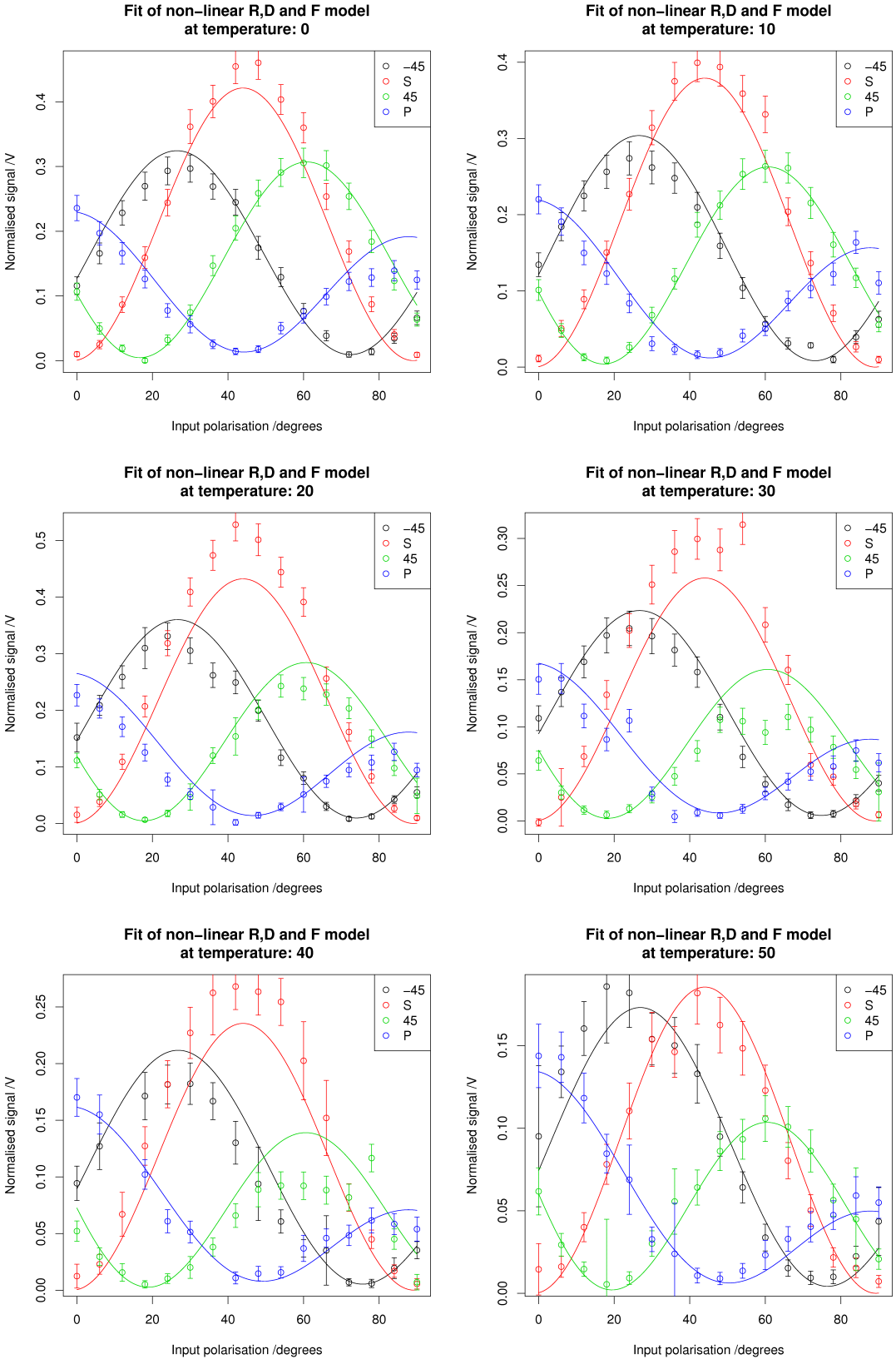


FIGURE 9.14: The summary plot of fitted curves against experimental data for each temperature from the 7.5 mM NB15C5 dataset fitted to the F, D and R model with fixed D, and $Im(R)$ and changing $Re(R)$. Water bath temperatures are given

As shown in table 9.12, the average values determined for D and $Im(R)$ are equal for both models. This suggests that these are likely to be a good estimate of the parameters, irrelevant of the model chosen. The errors in the regression results where $Re(R)$ is changing are marginally smaller than when fixed, showing this model to be more accurate. These changes in $Re(R)$ over the temperature range are small and only have a minor effect on the fitting, as shown in the two summary plots.

The results show that for either fitting assumption (fixed values or changing Re(R)) the surface orientation parameter, D, is fixed to 0.95 when using the refractive indices determined through ellipsometry and 0.77 when using the average of the refractive index of bulk water and air. These results correspond to an average molecular tilt angle, θ, of 12.9^o and 28.7^o respectively. The variation in these results show the significance in accurately determining the refractive indices. The fitting results shown in table 9.9 show both a smaller error and imaginary component in determining D when assuming ψ, the rotation about the main molecular axis, is random. The results following this have therefore assumed ψ to have a random distribution throughout. Fig. 9.15 shows a representation of how the determined angles correspond to the molecules at the surface. As in Fig. 9.15, the crown ring has not been included as the SHG results can only determine the orientation of the chromophore section of the molecule.

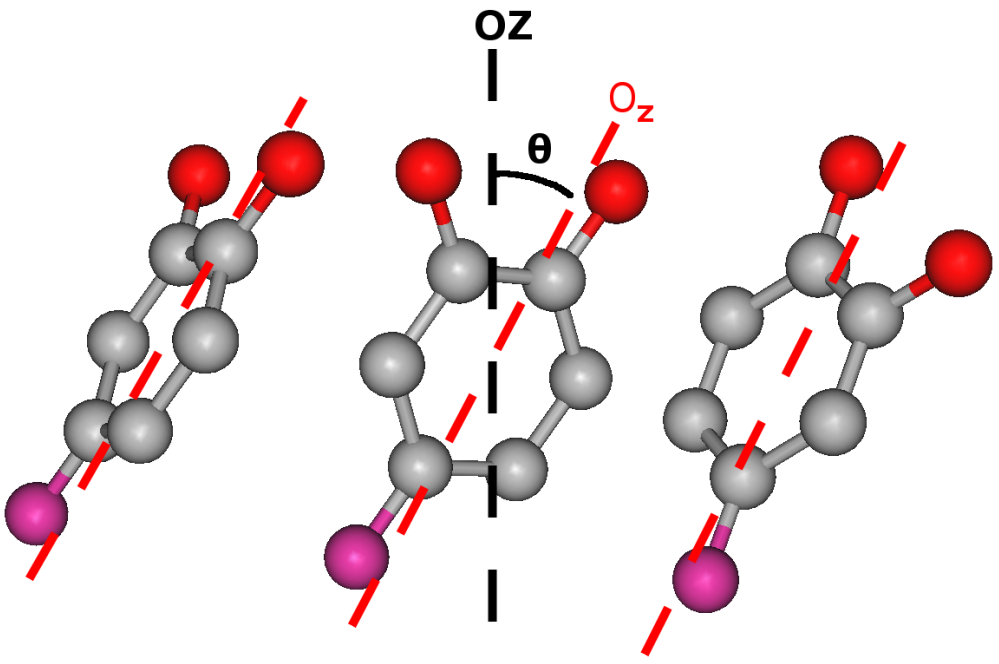


FIGURE 9.15: A representation of the alignment of the NB15C5 molecules at the surface based on the angles determined through the SHG study

Previous studies on crown ethers agree with the results collected in this study; showing a preferred orientation in the presence of a cyclic group on the ether ring to be tilting out of the water[125]. An SHG study of B15C5 calculated the molecular tilt angle, θ, to be 35.4-38.2^o [126]; showing only a small deviation from that of NB15C5. X-ray diffraction and Raman spectroscopic studies have shown crown ether compound to align in a rigid structure with bridging cations or infinite π stacking in the presence of a hydrogen bonding ligand[127, 128, 129]. The results from this study show that without the bridging cation the two dimensional rigidity is lost and the molecules are free to rotate about the molecular axis, ψ. Niga *et al* discuss a high value for the area per molecule for both NB15C5 and B15C5[130]; this supports the suggestion that the bridging between molecules is lost at the air/water interface and the rotation about the molecular axis is not fixed. This work is extended to look at the effect of concentration on the surface orientation[131]. These results show a significant shift in orientation with concentration. These effects have been studied further by SHG and are discussed further below.

9.3.2 Salt effect on surface orientation

Previous results by Timson *et al.*[104] suggested the addition of a salt affected the observed SHG signal and as such the surface orientation. These experiments were repeated with the aim to validate these findings and extend the research using lithium, sodium and potassium salts. In this previous work a manganese (II) salt was also used, although it was suggested a sodium impurity was present. As it is believed the results from the Mn(II) salt was due to this sodium impurity, manganese(II) was not included in this study.

Table 9.13 shows the ratio between fitted values for the salt data. These results have also been plotted, shown in Fig. 9.16.

Molar equiv of salt	B/A	C/A	B/C
1x LiCl	0.759±0.075	1.449±0.037	0.524±0.036
10x LiCl	0.758±0.129	1.389±0.065	0.545±0.067
1x NaCl	0.801±0.076	1.410±0.040	0.568±0.038
10x NaCl	0.678±0.079	1.192±0.042	0.569±0.055
1x KCl	0.730±0.077	1.315±0.039	0.555±0.045
10x KCl	0.943±0.055	1.473±0.033	0.640±0.026
None	0.774±0.050	1.530±0.023	0.506±0.022

TABLE 9.13: Non-linear fitting values determined from the 7.5 mM NB15C5 salt dependence dataset after applying a +3^o offset to the output polarisation

It can be seen from these results there has been very little effect on the surface orientation with the addition of any of the salts. It would have been expected that the sodium salt

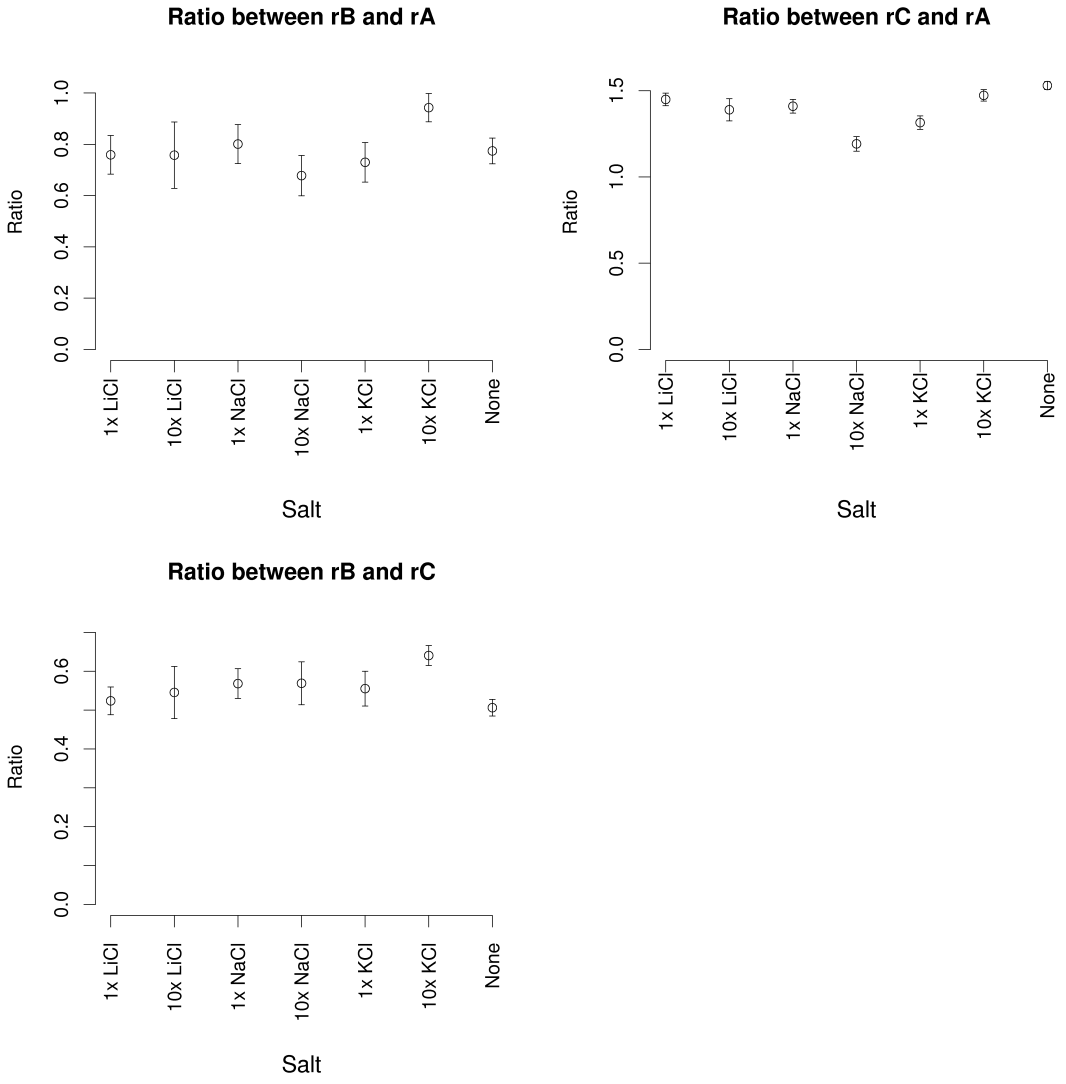


FIGURE 9.16: The plot of ratios between fitted values for the 7.5 mM NB15C5 salt dependence datasets

would have had the largest effect, with its ionic diameter similar to that of the crown ring diameter. It has been suggested potassium may also form a sandwich complex. It is believed the results seen in the previous work may be due to an impurity rather than the NB15C5 molecule. This was not seen in this work due to the extended purification process carried out before undertaking the experiments.

9.3.3 Adsorption isotherms

Fig. 9.17 shows a combined plot of both polarisations from the adsorption isotherm results. This plot shows that as the concentration changes, the ratio between the two polarisations changes. This suggests a change in surface orientation with respect to the

concentration of the bulk. This transition occurs at approximately 4.4mM. This effect has been studied further by a number of subsidiary experiments.

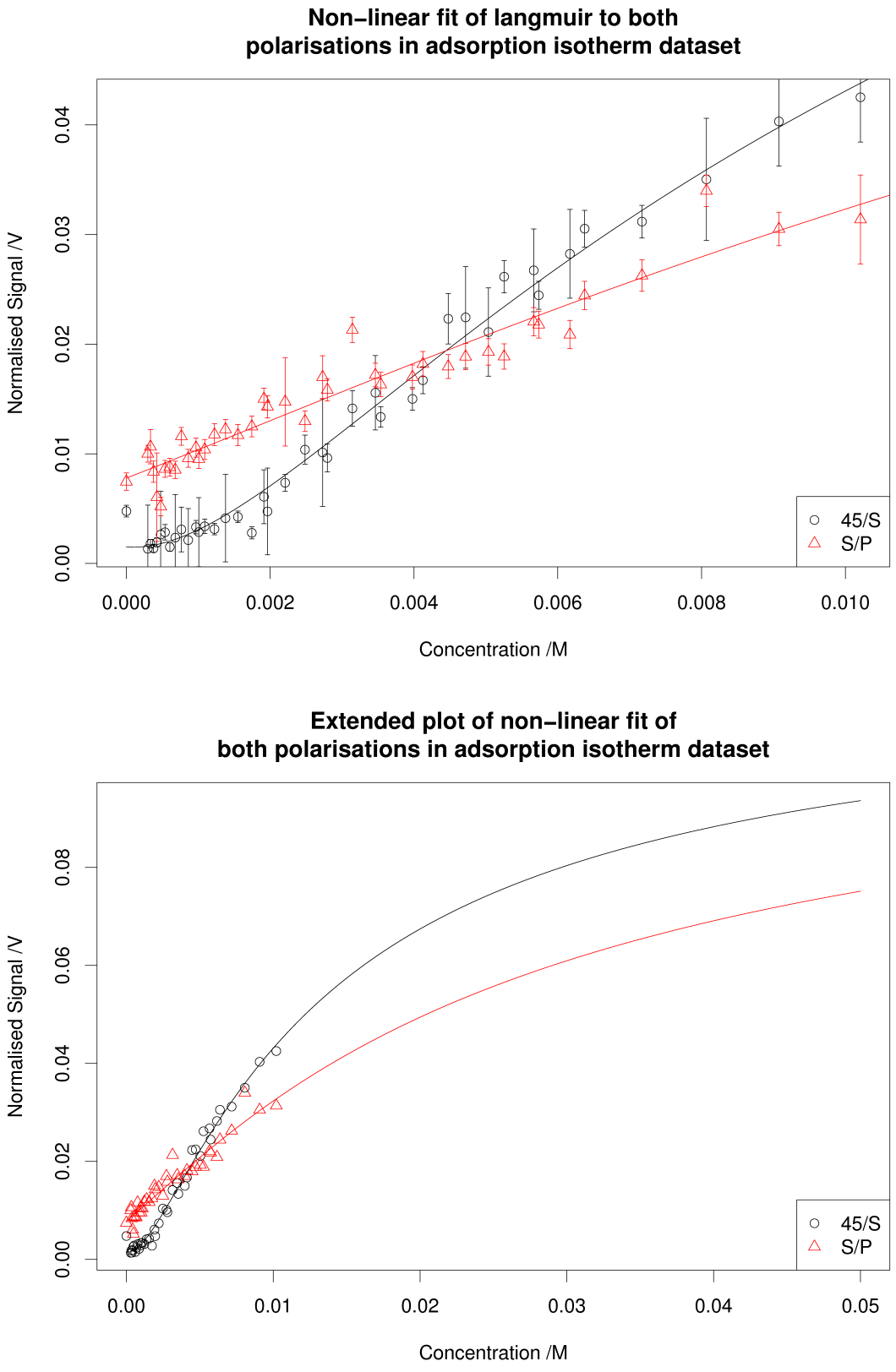


FIGURE 9.17: The combined plot of adsorption isotherm polarisation data for NB15C5 at 20^{o}C

9.3.3.1 Polarisation scans

A number of polarisation scan SHG experiments were carried out for different concentrations of the NB15C5 at 20^{o}C . This would allow a better understanding of how the orientation of the molecules is changing with respect to the concentration. The concentrations used in these experiments were 1.00, 3.04, 4.51 and 7.58 mM. The results from these experiments were fitted as with the previous polarisation scan experiments. The results from this fitting is shown in table 9.14 and Fig. 9.18. A plot of the fit to pure water has been included for reference. The ratios have been plotted in Fig. 9.19.

Concentration /mM	B/A	C/A	B/C
0.00	0.312±0.024	0.597±0.017	0.522±0.067
1.00	0.325±0.021	0.332±0.015	0.978±0.187
3.04	0.679±0.086	0.931±0.051	0.729±0.099
4.51	0.859±0.032	0.985±0.024	0.871±0.033
7.58	0.747±0.087	1.446±0.037	0.517±0.042

TABLE 9.14: Non-linear fitting values determined from the NB15C5 concentration dependence on surface orientation at 20^{o}C dataset after applying a +3^o offset to the output polarisation

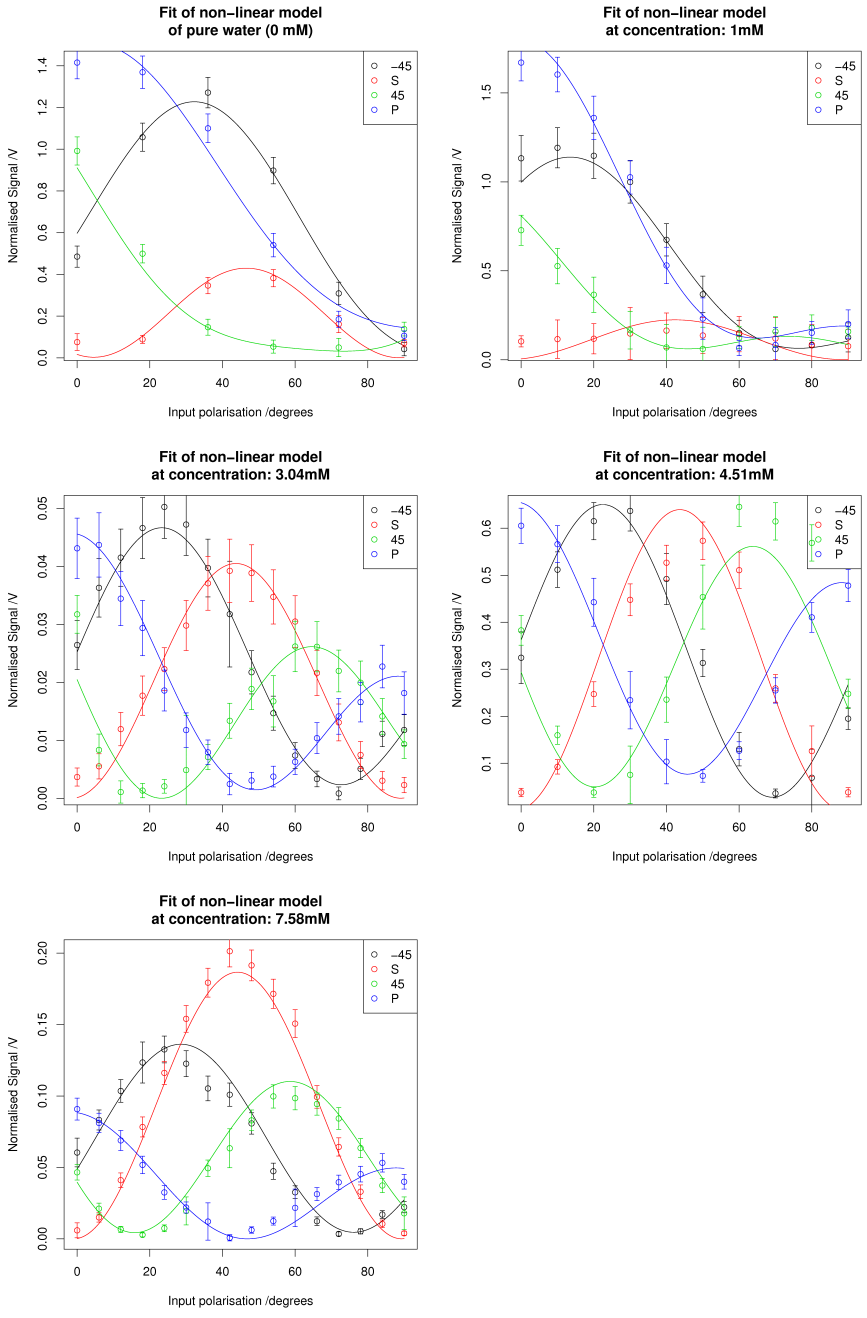


FIGURE 9.18: The summary plot of fitted curves for the concentration dependence on surface orientation data for NB15C5 at 20^{o}C . Error bars are $\pm 3\sigma$

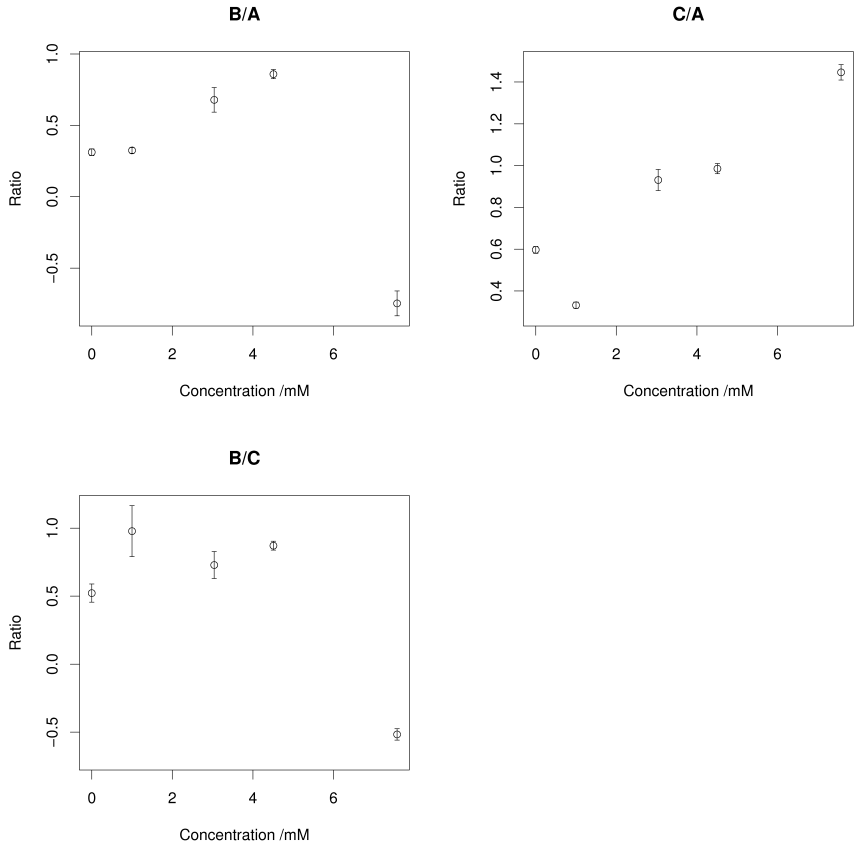


FIGURE 9.19: The plot of ratios between fitted values for the NB15C5 concentration dependence polarisation scan experiments

It can be seen from these results that there is a significant change in the ratios with concentration, and therefore a change in surface parameters. This can be seen when reviewing Fig. 9.18; the shape of the plots have changed dramatically over the different concentrations.

As with the previous results these were fitted to Eq. 6.28 and 6.29/6.30 to determine the ratio between dominant hyperpolarisability terms and the surface orientation parameter, D. These fitting were carried out using the refractive indices determined through the ellipsometry experiment, discussed above. These results can be seen in table 9.15, 9.16 and Fig. 9.20.

Conc. /mM	χ_{xxz}	χ_{zxx}	χ_{zzz}
0.00	7.547(±0.733) + 3.476(±1.242)i	7.020(±0.915) - 0.684(±4.004)i	309.856(±15.643) + 57.125(±19.643)i
1.00	4.820(±0.469) - 1.517(±0.721)i	-2.993(±0.817) + 7.454(±0.756)i	362.398(±8.523) - 75.223(±13.656)i
3.04	2.254(±0.036) + 0.097(±0.421)i	-2.376(±0.177) + 1.222(±0.346)i	93.380(±1.017) - 7.135(±8.500)i
4.51	8.155(±0.440) + 3.914(±0.859)i	-9.398(±1.157) - 8.703(±1.306)i	350.803(±11.861) + 120.082(±18.957)i
7.58	4.506(±0.207) - 1.868(±0.479)i	-4.079(±0.167) - 0.366(±0.471)i	156.015(±3.844) - 25.543(±9.063)i

TABLE 9.15: The molecular susceptibilities determined from the concentration dependence NB15C5 datasets carried out at 20°C

Conc. /mM	$\frac{\beta_{zxx}}{\beta_{zzz}}$	D ψ=90°	D ψ random
0.00	-0.008(±0.011) - 0.024(±0.017)i	0.951(±0.001) - 0.011(±0.009)i	0.951(±0.002) - 0.011(±0.009)i
1.00	-0.050(±0.003) + 0.038(±0.005)i	0.974(±0.001) + 0.002(±0.003)i	0.975(±0.001) + 0.002(±0.003)i
3.04	-0.096(±0.002) + 0.016(±0.006)i	0.956(±0.001) - 0.005(±0.004)i	0.958(±0.001) - 0.005(±0.004)i
4.51	-0.107(±0.002) - 0.032(±0.006)i	0.956(±0.001) - 0.004(±0.003)i	0.958(±0.001) - 0.004(±0.002)i
7.58	-0.104(±0.002) + 0.000(±0.004)i	0.946(±0.001) + 0.012(±0.003)i	0.949(±0.001) + 0.011(±0.003)i

TABLE 9.16: The molecular orientation parameters determined from the concentration dependence NB15C5 datasets carried out at 20°C

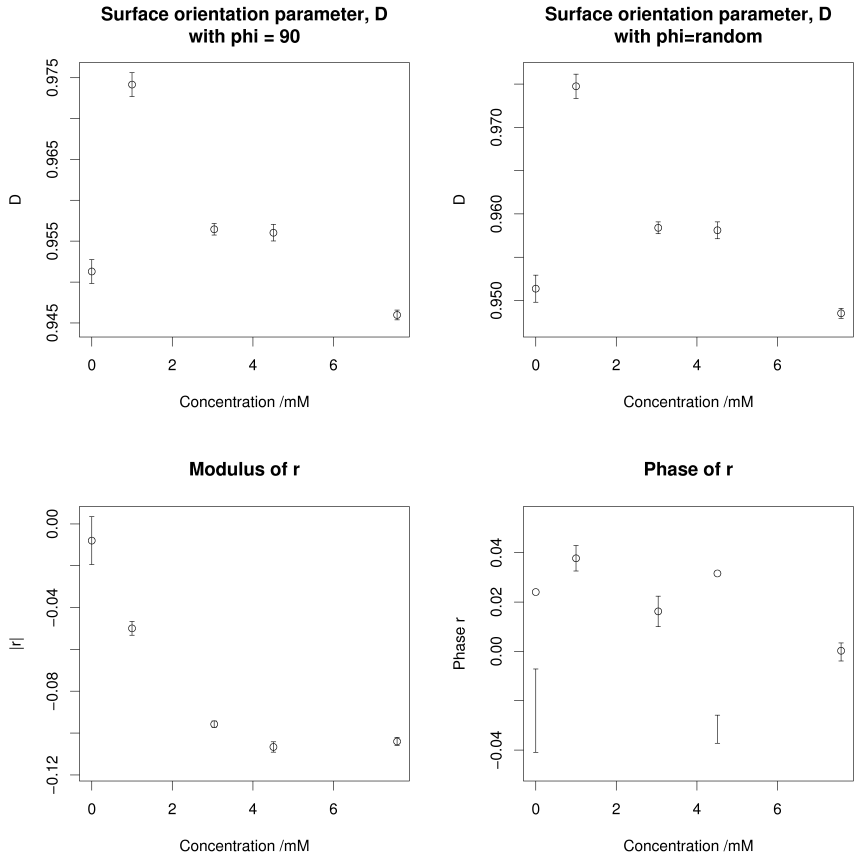


FIGURE 9.20: The plot of surface orientation parameter, D, and ratio between dominant hyperpolarisability terms for the NB15C5 concentration dependence datasets completed at 20^{o}C

As with the temperature dependence results, there is a change in the ratio between the dominant hyperpolarisability terms with changes in concentration. In the case of using the ellipsometry refractive indices this is dominated by the β_{zzz} term, although when using average refractive indices β_{zxx} must also be considered. This suggests a change in the electrical configuration of the molecule as the bulk concentration is increased.

9.3.3.2 Surface tension

The surface tension experiments were carried out using a torsion balance as described in section 7.2.1. The auto-diluter used in the SHG experiments was used to sequentially dilute the sample after each set of surface tension measurements had been made. A 10mM NB15C5 solution was prepared and underwent 26 sequential 3ml dilutions of water. At each concentration a minimum of three surface tension measurements were made and the average calculated. The measurements were carried out at 20^{o}C . The results are presented in table 9.17, Fig. 9.21 shows a plot of this data. The SHG adsorption isotherm results have been included on a secondary axis to allow comparison.

		Measurement				
Concentration /mM	Average /mNm^{-1}	1st	2nd	3rd	4th	5th
10.0	52.88	52.5	53.0	53.0	53.0	
9.4	54.17	54.0	54.5	54.0		
8.8	55.00	55.0	55.0	55.0		
8.3	56.20	56.0	56.0	55.0	56.0	58.0
7.8	57.40	58.4	56.7	57.5	57.0	
7.3	60.75	61.0	60.0	61.0	61.0	
6.9	59.63	60.0	59.0	60.0	59.5	
6.5	61.50	59.5	60.0	62.5	62.5	63.0
6.1	65.67	65.0	66.0	66.0		
5.7	67.17	67.0	67.0	67.5		
5.4	68.30	68.0	68.5	68.4		
5.1	69.33	69.5	69.0	69.5		
4.8	69.83	70.0	69.5	70.0		
4.5	70.10	70.0	70.0	70.3		
4.2	70.13	70.0	70.4	70.0		
3.8	69.30	69.5	69.4	69.0		
3.4	69.83	70.0	69.5	70.0		
3.0	69.47	69.5	69.5	69.4		
2.5	70.53	70.5	70.6	70.5		
2.0	71.00	71.0	71.0	71.0		
1.6	71.93	72.3	71.5	72.0		
1.3	72.00	72.0	72.0	72.0		
1.0	71.83	72.0	71.5	72.0		
0.7	71.60	71.8	71.0	72.0		
0.5	71.83	72.0	72.0	71.5		
0.3	72.77	72.8	72.5	73.0		
0 (pure water)	73.00	73.0	73.0	73.0		

TABLE 9.17: Surface tension results from the NB15C5 adsorption isotherm at 20^{o}C

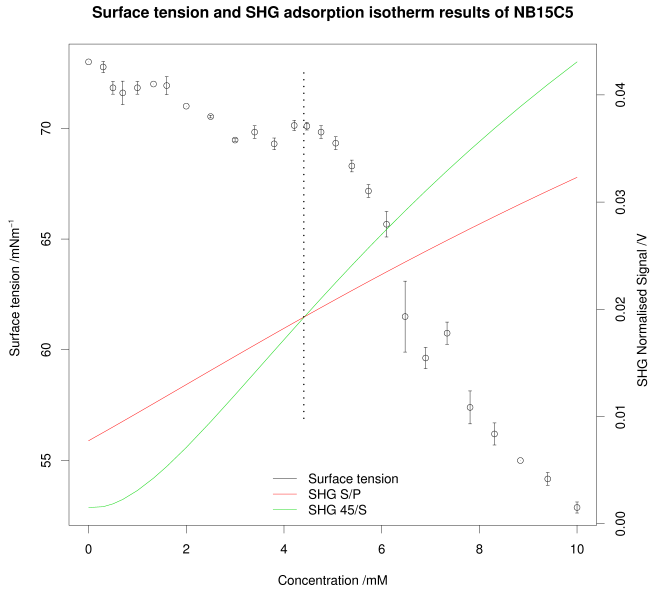


FIGURE 9.21: Plot of surface tension and SHG results from the NB15C5 adsorption isotherm at 20^{o}C

These results show an unexpected change in surface tension at a concentration of 4.4 mM, as with the SHG adsorption isotherm experiment. This supports the theory that a surface orientation change is occurring at this concentration.

9.3.3.3 Ellipsometry

As discussed in chapter 1, ellipsometry is used to study films at the interface. In this investigation spectroscopic ellipsometry has been used. This method scans through a range of wavelength to determine changes in properties with respect to wavelength. The instrument used in these experiments was a UVISEL VUV phase modulated spectroscopic ellipsometer from Horiba Jobin Yvon. A 7.5 mM NB15C5 solution was prepared and aligned within the instrument. A VAS experiment was completed to determine the optimal angle of incidence. The instrument was set up to run from 50^o to 60^o angle of incidence in steps of 2^o . From this experiment the optimal angle was found to be 54^o and as such the instrument set to this for further experiments. A spectroscopic experiment was carried out for a number of concentrations of NB15C5, scanning the wavelength from 300 nm to 700 nm in 5 nm steps. The concentrations used in the experiment were 7.53, 6.1, 4, 2.92, 1.91 and 1.02 mM. Fig. 9.22 shows a plot of delta and Fig. 9.23 a plot of psi for these concentrations.

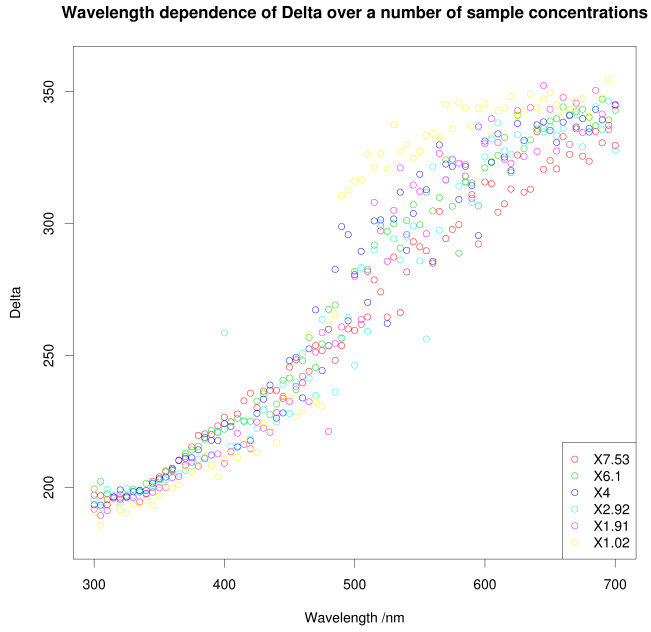


FIGURE 9.22: A plot of delta with respect to wavelength for a number of NB15C5 sample concentrations determined via spectroscopic ellipsometry

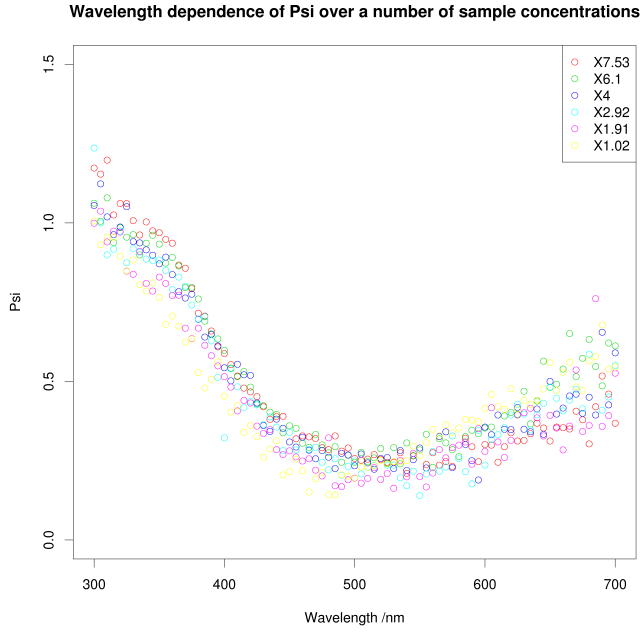


FIGURE 9.23: A plot of psi with respect to wavelength for a number of NB15C5 sample concentrations determined via spectroscopic ellipsometry

These results show a shoulder in the ψ curve at approximately 350 nm, corresponding to the adsorption peaks seen in the UV/VIS shown in Fig. 9.24.

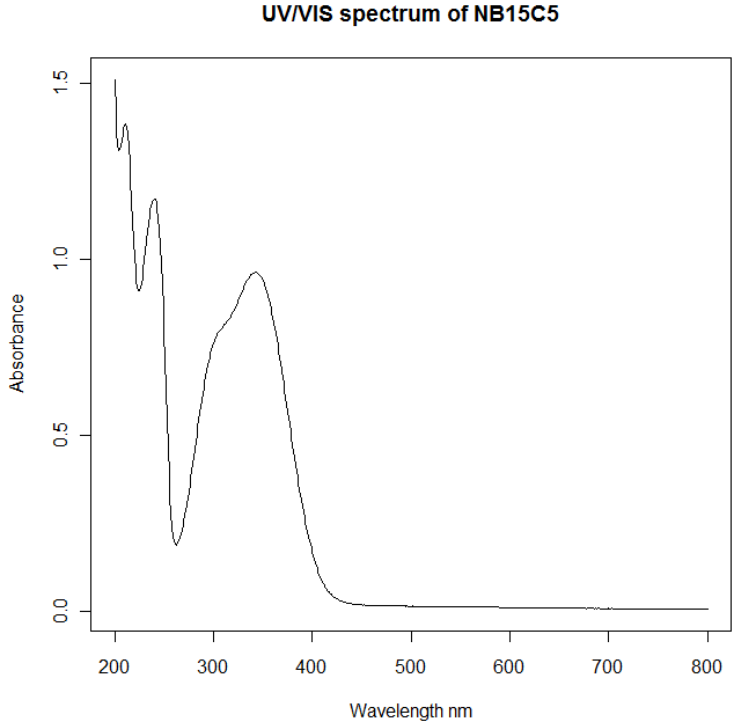


FIGURE 9.24: UV/Vis spectrum of a 0.75 mM NB15C5 solution

As with the previous ellipsometry data, this was fitted to a Quatre Oscillator using a three oscillator model; accounting for the dense layer of water at the surface and the two oscillators observed in the UV/VIS of NB15C5. The oscillator values were seeded with those obtained from the UV/VIS spectrum. When fitting these results the software failed to converge on reasonable values, deviating significantly from the seeded values. The layer thickness determined from this fitting also containing a large error. It is thought this fitting failed as the parameters relating to the two NB15C5 oscillators will be correlated and therefore complicate the fitting process. The results determined through this work will be investigated in further studies.

It was hoped to determine through this study how the surface thickness and refractive indices are affected by the concentration dependent phase transition observed in adsorption isotherms. A change in surface thickness may signify a change in surface orientation not observed in the SHG experiments. It is expected the refractive indices would tend towards those values determined for the dense layer of water. These may have been linear with concentration or show a similar transition to that observed in the SHG experiments at a concentration of 4.4 mM.

9.3.4 Further work

As shown in the adsorption isotherm results, there is a phase transition occurring with change in concentration at 4.4 mM. The temperature dependence on surface orientation results were completed with a concentration of 7.5 mM and as such only includes information on the higher phase. A repeat of these temperature dependence experiments can be carried out at the lower phase and transition point to extend the understanding of the changes occurring. To complete these experiments a high power laser source would be required to obtain reliable data. The adsorption isotherm experiments can also be repeated at multiple temperatures to determine the effect temperature has on the phase transition and adsorption at the interface.

Chapter 10

Benzo-15-crown-5

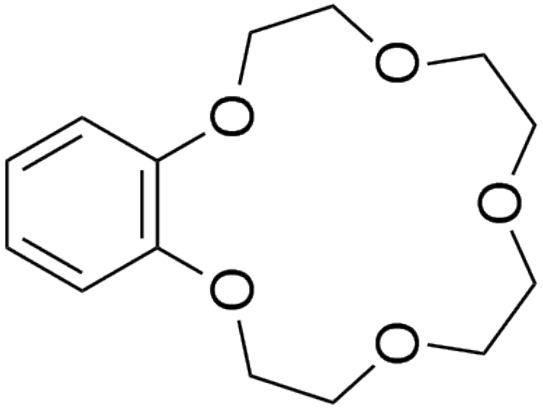


FIGURE 10.1: The molecular structure of B15C5

As discussed in the previous chapter, crown ethers are important due to their selective binding to ions and in facilitating the transfer of ions between inorganic and organic solvents. As a way of extending the study of this group of compounds, B15C5 was included. As these compounds have such importance in binding with ions, the effect of addition of a salt will be investigated. Before these results can be fully understood the free-form compound must first be investigated.

10.1 Experimental

As discussed by Rousay *et al.*[96], an off-the-shelf sample of B15C5 (Fluka >99.0% purity) was found to contain impurities which will affect the surface of the solution. It is believed sodium is leached from the glassware and as such the 'pure' compound already contains a certain amount of complexed sample. To overcome this, the bought sample was purified before use.

The solid was dissolved in excess heptane and heated to 70^oC, A dark brown oil was observed in the bottom of the vessel. The solution was vacuum filtered hot and the filtrate collected. This process of heating to 70^oC and filtering was repeated at least three times. The heptane was removed under vacuum to produce very fine, clear crystals. The purified compound was stored in glass bottles.

In the cited work a comparison between the purified compound stored in glass and Teflon was completed. From this study it was found that the compound stored in glass showed similar properties to those when excess sodium chloride was added, suggesting the compound was leaching sodium contained within the glass vial. This comparison was repeated via SHG polarisation scans to verify these results. A sample of B15C5 was purified using the process described above and dried on a high vacuum line for approximately 15 minutes. The compound was collected and dissolved in pure water to produce a 6 mM solution. A full polarisation scan was completed for this sample. From starting the purification to completing the experiment was under four hours, this should minimise any sodium leaching should it be occurring. A second solution using compound which had been stored in the glass bottle for over a month underwent the same procedure. Fig 10.2 shows the plots of the data from these two experiments.

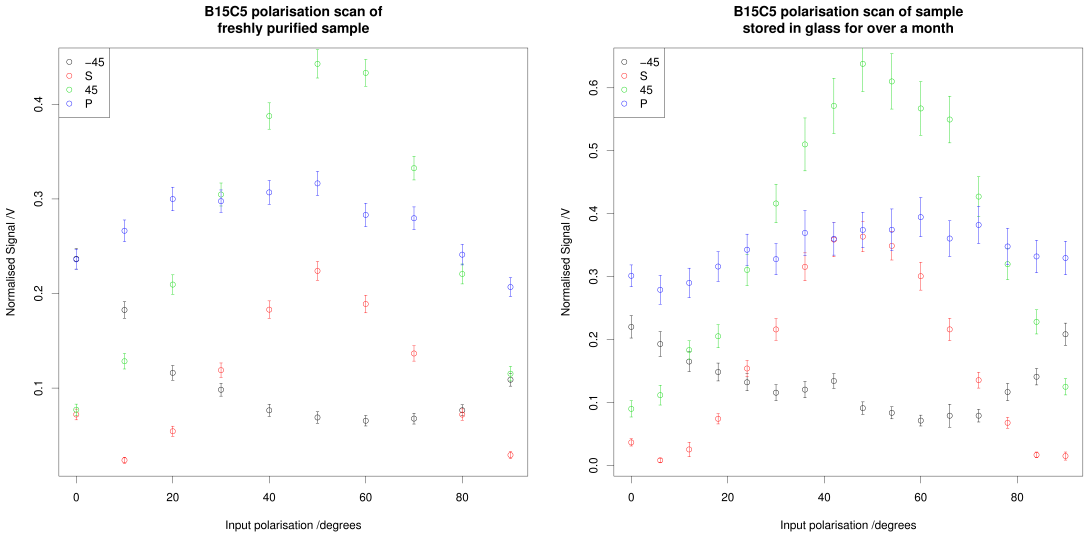


FIGURE 10.2: Polarisation scan results from two B15C5 samples. Left showing freshly purified sample and right stored in glass for over a month

It can be seen from these results there is very little change between the two datasets. The difference between the normalised intensities results from variations in laser power and positioning of the photodiode during the experiments (as there is a 5 month gap between the two experiments). Therefore all B15C5 samples will be considered in their free state unless a salt has been specifically added. It is thought the impurities seen in the previous work was due to the experimental practises carried out; the purified compound still containing a number of impurities.

After purification, it was possible to produce a solution in water with a maximum concentration of approximately 30 mM. Many experiments in this study use a concentration of 7.50 mM; to prepare this samples the solutions were heated to 60^{o}C and sonicated while hot for approximately two hours. For concentrations higher than this the sample required longer sonication.

10.1.1 Temperature dependence of surface orientation

As with the previous samples, a study of the temperature dependence of surface orientation was conducted. The study investigated orientation changes which occur between 0^{o}C and 50^{o}C . These experiments were conducted using the water bath described in section 6.4.6 and followed the work-flow described in section 6.5.2.1. These experiments were run at water bath temperatures between 0^{o}C and 50^{o}C in 10^{o}C steps. A 7.50 mM B15C5 solution was used throughout, although a fresh sample was prepared each day.

10.1.2 Effects of salt on surface orientation

As the results of the NB15C5 salt experiments showed no effect on surface orientation the same experiments were repeated with the B15C5. It was hoped to see no effect; this would validate the theory that the previous results were caused by an impurity rather than the sample as the compounds are structurally similar. These experiments were run by addition of the chloride salt of lithium, sodium and potassium during the sample preparation stage. Each experiment was run at both a 1x and 10x molar equivalent addition of the salt. A 7.50 mM solution was used throughout and the waterbath maintained at 20^{o}C . The workflow described in section 6.5.2.1 was followed to complete these experiments.

10.1.3 Adsorption isotherm

An adsorption isotherm of B15C5 was completed to determine the surface equilibrium constant, k, at 20^oC. A 50 ml solution of 7.50 mM B15C5 underwent a number of sequential dilutions until the signal had drop to that of water. At each concentration the SHG signal was captured over 500 laser shots for the polarisation combination of 45/S. A second solution of pure water was concentrated by a number of sequential additions to a concentration of approximately 7.5 mM.

10.2 Results

10.2.1 Temperature dependence of surface orientation

Table 10.1 describes the combination of experiments used to make the B15C5 temperature dependence dataset. As with the previous results there were issues with evaporation at higher temperature and as such these are made up of a number of shorter experiments.

Water bath temp. /^{o}C	Experiment number	Experimental data
0	450	S output
		P output
		45 output
		-45 output
10	451	S output
		P output
		45 output
		-45 output
20	443	S output
		P output
		45 output
		-45 output
30	444	S output
		P output
		45 output
		-45 output
40	445	S output
		45 output
	448	P output
	447	-45 output
50	453	S output
	454	45 output
	455	P output
	456	-45 output

TABLE 10.1: The combination of experiments used to generate the temperature dependence of surface orientation dataset for B15C5. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

This dataset was fitted via non-linear regression to Eq. 6.21 using an R script. The model was set up to consider A real with B and C complex. A plot of the 20^{o}C data is shown in Fig. 10.3. It can be seen from this fitting the standard equation is unsuitable for the experimental results.

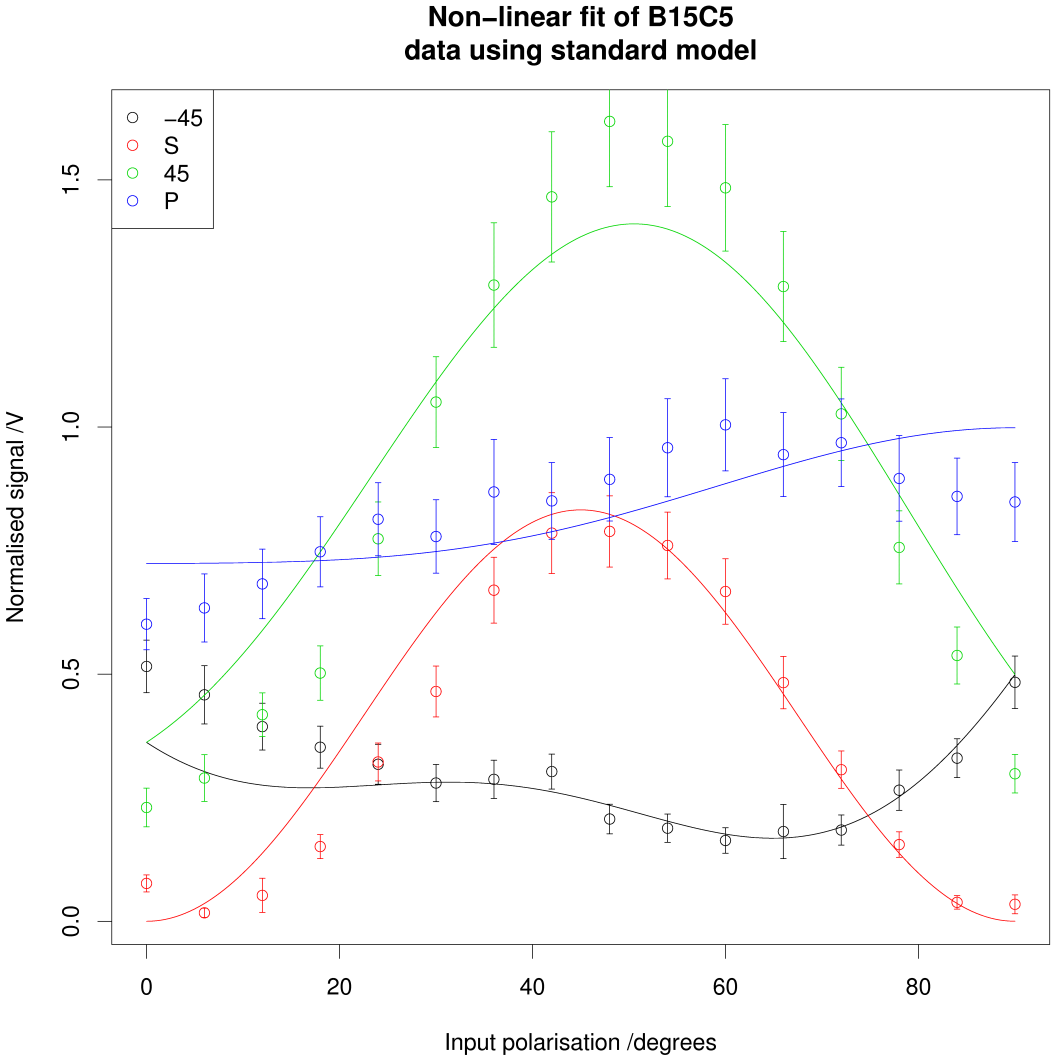


FIGURE 10.3: Plot of fitted values to the 20^{o}C B15C5 experimental data, showing the poor fitting using the original fitting equation

The equation can be modified to include two extra terms, D and E, which account for chirality in the system. The reasoning for including chiral terms with an achiral compound will be considered in the discussion section of this chapter. As each of these new terms can be either real or complex an ANOVA study was carried out on the 20^{o}C data. This will determine if allowing either or both parameters to be complex will have a significant improvement of the fitting of the data. The ANOVA results are shown in table 10.2 and Fig. 10.4 shows a comparison between the basic and new fitting procedures, showing the improvement on fitting to the data.

Model	A	B	C	D
ADE real	0.766±0.014	0.865(±0.031) - 0.260(±0.093)i	-0.697(±0.043) + 0.702(±0.030)i	0.140±0.018
AD real	0.772±0.014	0.830(±0.043) - 0.354(±0.094)i	-0.636(±0.051) + 0.786(±0.045)i	0.146(±0.019)
AE real	0.791±0.029	-0.872(±0.066) + 0.216(±0.194)i	0.414(±0.101) - 0.752(±0.051)i	0.069(±0.060) + 0.850(±0.079)i
A real	0.816±0.014	0.820(±0.030) + 0.351(±0.053)i	-0.314(±0.064) - 0.957(±0.024)i	-0.013(±0.039) + 0.359(±0.054)i

Model	E	RSE
ADE real	0.286±0.023	0.046
AD real	0.254(±0.019) - 0.094(±0.046)i	0.045
AE real	0.021±0.041	0.071
A real	0.190(±0.025) + 0.172(±0.028)i	0.037

	ANOVA	
Model	F test	Pr(>F)
ADE real		
AD real	3.904	0.053
AE real	-31.746	1.000
A real	27.897	2.251e-06

TABLE 10.2: Non-linear fitted values determined from the 20^{o}C B15C5 data, comparing the results by varying which factors are considered complex. B and C are considered complex throughout. (RSE - Residual standard error)

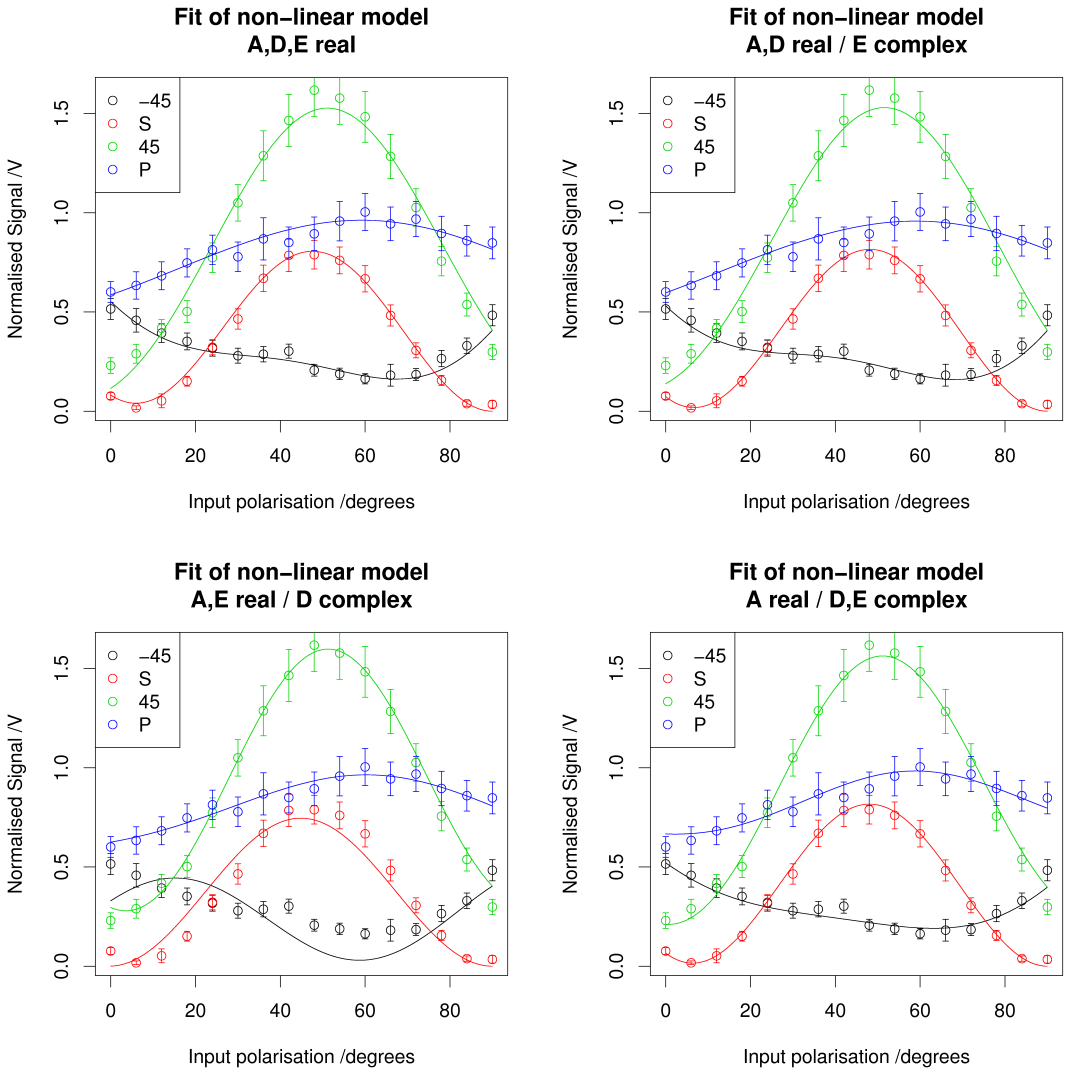


FIGURE 10.4: Plot of fitted values to the 20^{o}C B15C5 experimental data, comparing fits by varying which factors are considered complex

The ANOVA results show a significant improvement when considering both the D and E terms to be complex. Therefore, for subsequent fitting in this chapter the data will be fitted to that shown in Eq. 10.1, a modified version of the fitting equation used for the previous compounds, Eq. 6.21. This includes the chiral term of the susceptibility, χ_{XXZY}.

$$I(2\omega) = |(C\sin 2\gamma + E\cos^2\gamma)\sin\Gamma + (A\cos^2\gamma + B\sin^2\gamma + D\sin 2\gamma)\cos\Gamma|^2 \quad (10.1)$$

The temperature dependence results were fitted to this new equation. Table 10.3 shows the fitted values and Fig. 10.5 shows a summary plot of the fitted values against the experimental data.

Water bath temp.^{o}C	A	B	C	D
0	0.615±0.011	0.673(±0.018) - 0.157(±0.052)i	-0.418(±0.061) + 0.688(±0.037)i	0.054(±0.030) - 0.220(±0.059)i
10	0.644±0.011	0.643(±0.026) - 0.296(±0.047)i	-0.383(±0.056) + 0.784(±0.028)i	0.061(±0.028) - 0.196(±0.050)i
20	0.816±0.014	0.820(±0.030) + 0.351(±0.053)i	-0.314(±0.064) - 0.957(±0.024)i	-0.013(±0.039) + 0.359(±0.054)i
30	0.872±0.017	0.854(±0.039) - 0.365(±0.071)i	-0.386(±0.072) + 0.895(±0.033)i	0.029(±0.049) - 0.394(±0.070)i
40	0.817±0.032	0.820(±0.104) - 0.441(±0.183)i	-0.647(±0.149) + 0.842(±0.115)i	0.184(±0.064) - 0.050(±0.170)i
50	0.771±0.021	0.910(±0.046) - 0.386(±0.088)i	-0.412(±0.098) + 0.950(±0.044)i	0.063(±0.057) - 0.332(±0.087)i

Water bath temp.^{o}C	E	Residual standard error
0	0.184(±0.029) - 0.153(±0.019)i	0.023
10	0.168(±0.023) - 0.161(±0.018)i	0.024
20	0.190(±0.043) + 0.172(±0.021)i	0.037
30	0.192(±0.067) - 0.123(±0.025)i	0.046
40	0.202(±0.054) - 0.256(±0.059)i	0.090
50	0.127(±0.071) - 0.176(±0.031)i	0.053

TABLE 10.3: Non-linear fitting values determined from the B15C5 temperature dependence dataset. The extended fitting equation has been used

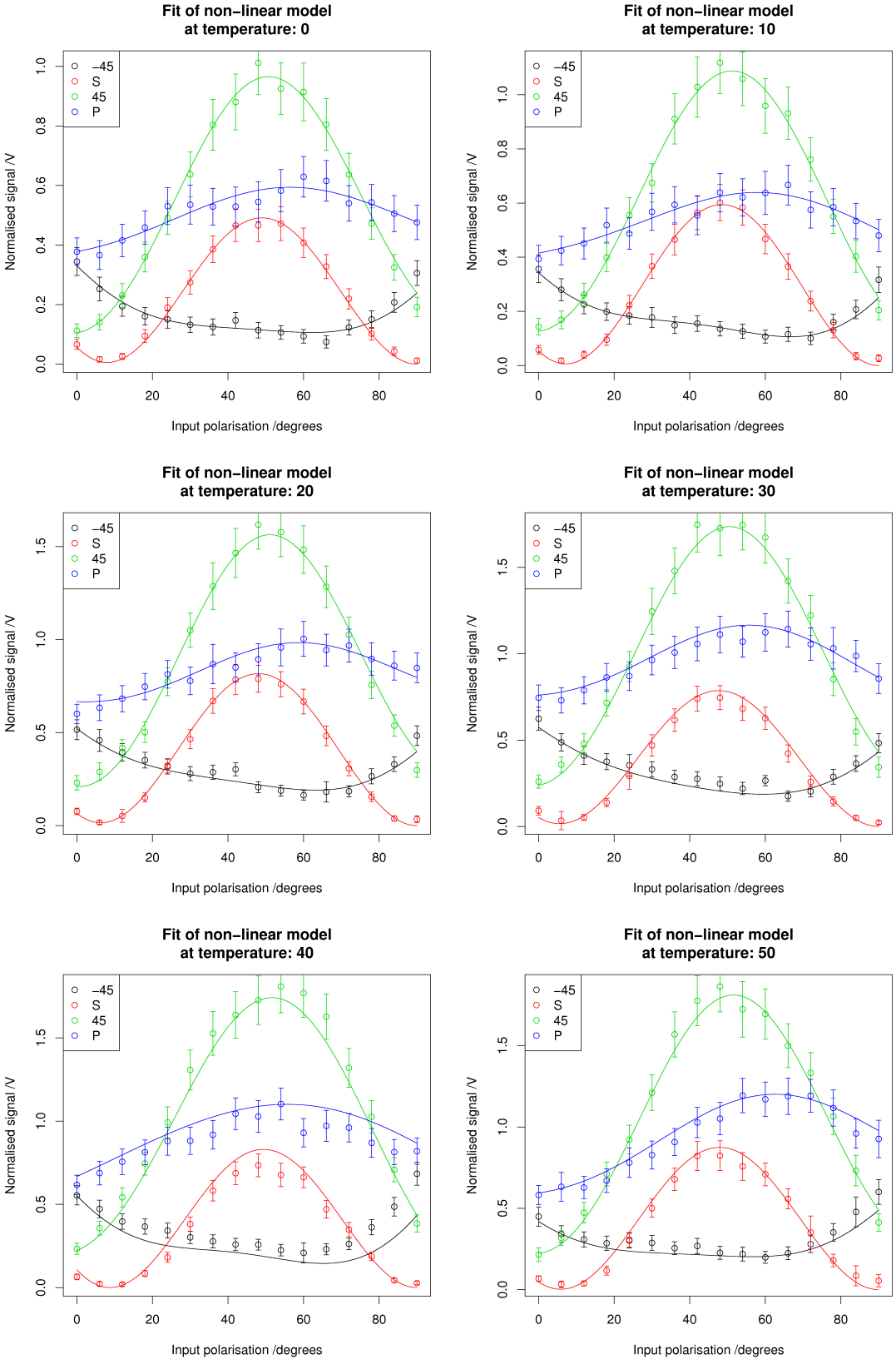


FIGURE 10.5: The summary plot of fitted curves against experimental data for each temperature from B15C5 temperature dependence dataset after applying the $+3^o$ offset to the output polarisation. The extended fitting equation has been used. Water bath temperatures are given. Error bars are $\pm 3\sigma$

10.2.2 Effects of salt on surface orientation

Table 10.4 describes the salt dependence dataset.

Salt equiv	Experiment number
1x LiCl	500
10x LiCl	501
1x NaCl	480
10x NaCl	481
1x KCl	482
10x KCl	483
None	443

TABLE 10.4: The combination of experiments used to generate the salt dependence of surface orientation dataset for B15C5. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

The results were fitted to the extended fitting equation, Eq. 10.1. Table 10.5 shows the fitted values and Fig. 10.6 shows a summary plot of fitted curves.

Salt	A	B	C	D
1x LiCl	0.165±0.004	0.161(±0.009) - 0.090(±0.014)i	-0.058(±0.016) + 0.195(±0.006)i	0.004(±0.009) - 0.060(±0.013)i
10x LiCl	0.182±0.005	0.197(±0.008) - 0.045(±0.023)i	-0.090(±0.030) + 0.206(±0.013)i	0.017(±0.013) - 0.059(±0.026)i
1x NaCl	0.564±0.031	0.437(±0.207) - 0.481(±0.180)i	-0.344(±0.116) + 0.614(±0.065)i	0.149(±0.087) - 0.195(±0.099)i
10x NaCl	0.616±0.018	0.519(±0.091) - 0.486(±0.090)i	-0.301(±0.064) + 0.678(±0.031)i	0.096(±0.050) - 0.223(±0.060)i
1x KCl	0.548±0.011	0.579(±0.019) - 0.195(±0.042)i	-0.282(±0.058) + 0.706(±0.024)i	0.034(±0.027) - 0.184(±0.048)i
10x KCl	0.481±0.021	0.407(±0.129) - 0.487(±0.102)i	-0.329(±0.078) + 0.661(±0.039)i	0.118(±0.059) - 0.200(±0.058)i
None	0.816±0.014	0.820(±0.030) + 0.351(±0.053)i	-0.314(±0.064) - 0.957(±0.024)i	-0.013(±0.039) + 0.359(±0.054)i

Salt	E	Residual standard error
1x LiCl	0.025(±0.009) - 0.039(±0.006)i	0.002
10x LiCl	0.042(±0.016) - 0.025(±0.008)i	0.003
1x NaCl	0.073(±0.061) - 0.166(±0.085)i	0.055
10x NaCl	0.066(±0.034) - 0.193(±0.042)i	0.035
1x KCl	0.156(±0.022) - 0.157(±0.017)i	0.020
10x KCl	0.038(±0.040) - 0.172(±0.046)i	0.031
None	0.190(±0.043) + 0.172(±0.021)i	0.037

TABLE 10.5: Non-linear fitting values determined from the B15C5 salt dependence dataset after applying a $+3^o$ offset to the output polarisation

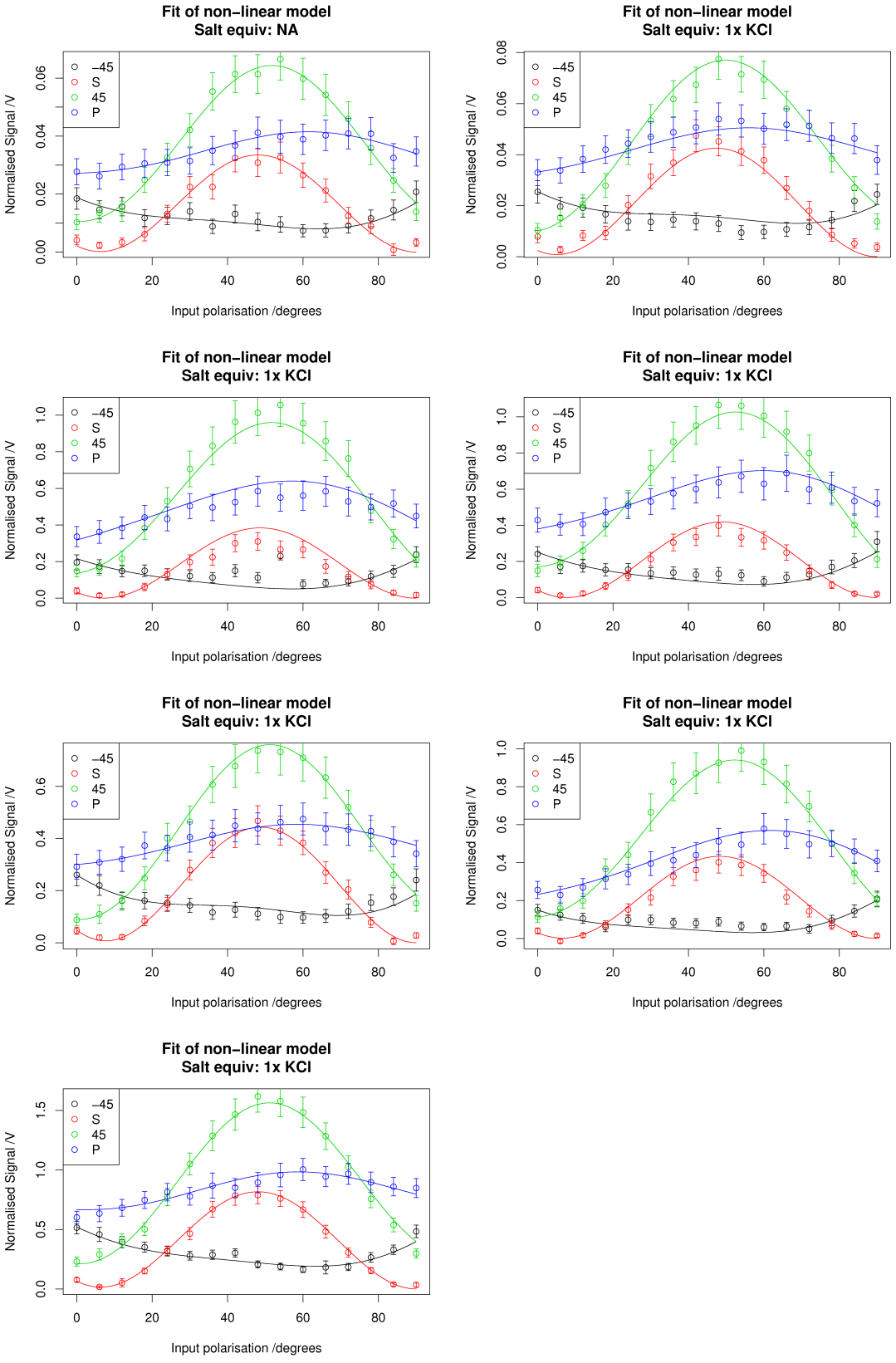


FIGURE 10.6: The summary plot of fitted curves against experimental data for the salt dependence B15C5 data after applying the $+3^o$ offset to the output polarisation. Error bars are $\pm 3\sigma$

10.2.3 Adsorption isotherm

After initial experiments running the adsorption isotherm of B15C5 it was found that compound was 'locking' onto the surface. When running the concentration from high to low the signal would not drop until a very low concentration. A description and possible cause of this phenomenon will be included in the discussion section of this chapter. After observing these results the experiment was modified in an attempt to capture all the information on the system. A solution of pure water first underwent a series of serial additions of a B15C5 stock solution until the signal levels had reached a plateau. A second experiment was completed with a 'fresh' 7.5 mM B15C5 solution which underwent a number of serial dilutions with pure water. The experiment was continued until the signal had dropped to that of which was expected from water. The combination of experiments completed to make up the two dataset are described in table 10.6. These experiments were run at a single polarisation combination of 45/S.

Experiment type	Experiment number	Concentration range /mM
Concentrating sample	228	0 - 0.41
	226	0.41 - 4.13
	227	4.13 - 17.06
Diluting sample	253	7.50 - 3.07
	254	3.07 - 0.04
	255	0.04 - 0.41×10^{-3}
	256	0.41×10^{-3} - 4.38×10^{-5}

TABLE 10.6: The combination of experiments used to generate the B15C5 adsorption isotherm datasets. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

Fig. 10.7 shows a plot at the experimental results, a second plot of lower concentrations has been included to show the drop in signal for the dilution isotherm. It can be seen that these results do not follow the Langmuir adsorption model assumptions and as such were not fitted as with the previous adsorption isotherm results, but are discussed below.

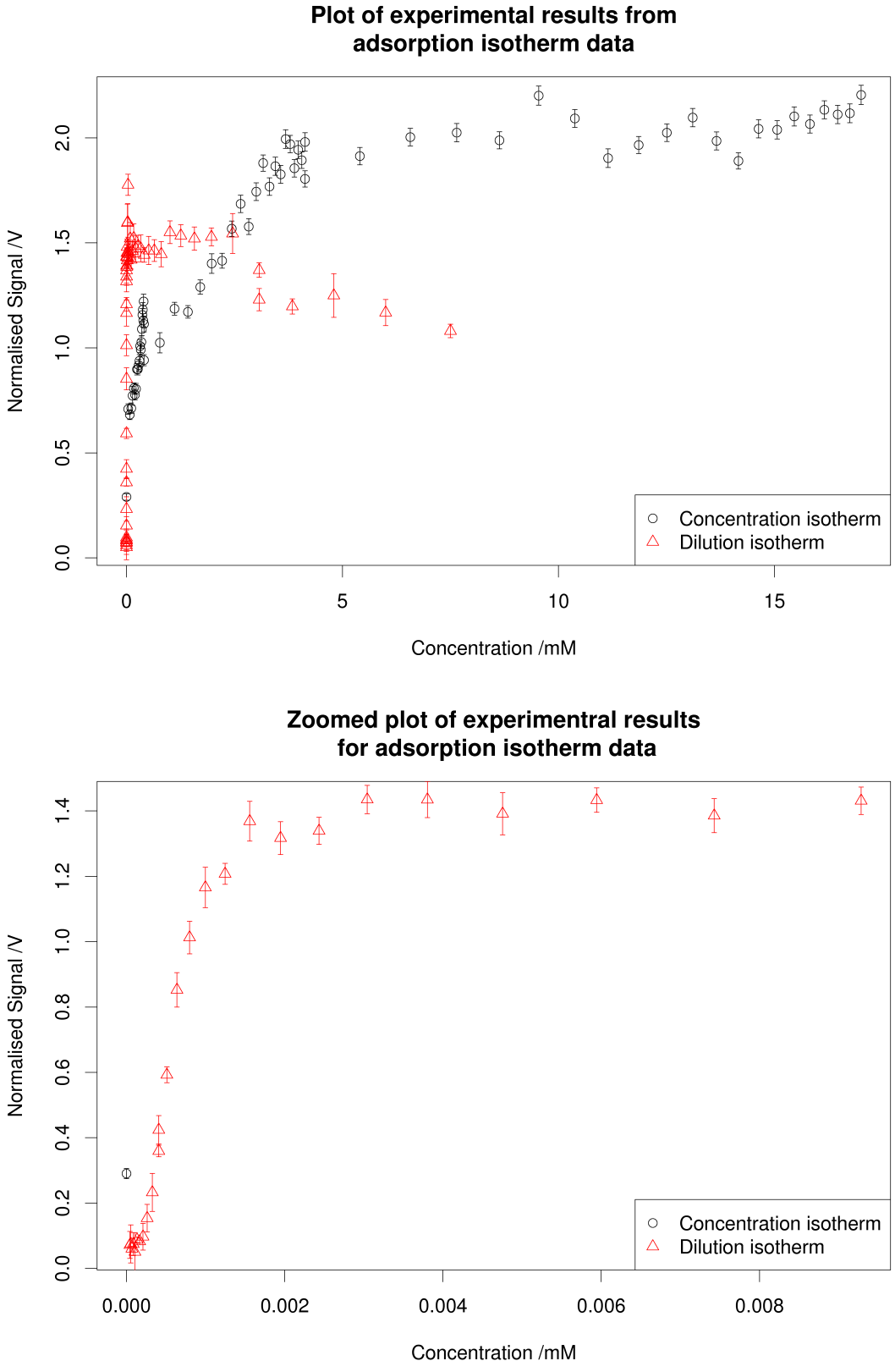


FIGURE 10.7: The plot of the experimental results from the B15C5 adsorption isotherm experiments at 20^{o}C . Top: Full results set. Bottom: Zoomed section showing drop in signal for dilution isotherm.

10.3 Discussion

10.3.1 Temperature dependence of surface orientation

Table 10.7 shows the ratios between the fitted parameters for the B15C5 temperature dependence results and the fitted curves are shown in Fig. 10.8. The temperatures reported in there results have been converted from water bath temperature to sample temperature.

These results show little trend with changes in temperature, therefore it is assumed there is little change in surface orientation with respect to temperature. The largest errors are observed in the higher temperature results, as expected due to the evaporation issues during the experiments.

Sample temp. /K	B/A	C/A	B/C	D/B	D/C	D/E
278.5	1.124±0.024	1.310±0.034	0.858±0.014	0.327±0.107	0.281±0.079	0.946±0.895
286.1	1.099±0.021	1.355±0.029	0.812±0.012	0.290±0.081	0.235±0.053	0.881±0.746
293.7	1.093±0.017	1.234±0.024	0.886±0.011	0.403±0.069	0.357±0.054	1.402±0.834
301.3	1.065±0.018	1.118±0.026	0.953±0.014	0.426±0.077	0.405±0.070	1.732±1.279
308.9	1.140±0.039	1.299±0.054	0.877±0.023	0.204±0.045	0.179±0.034	0.583±0.363
316.6	1.283±0.024	1.344±0.037	0.955±0.014	0.342±0.077	0.326±0.070	1.556±1.592

TABLE 10.7: The ratio between fitted values from the 7.5 mM B15C5 temperature dependence datasets

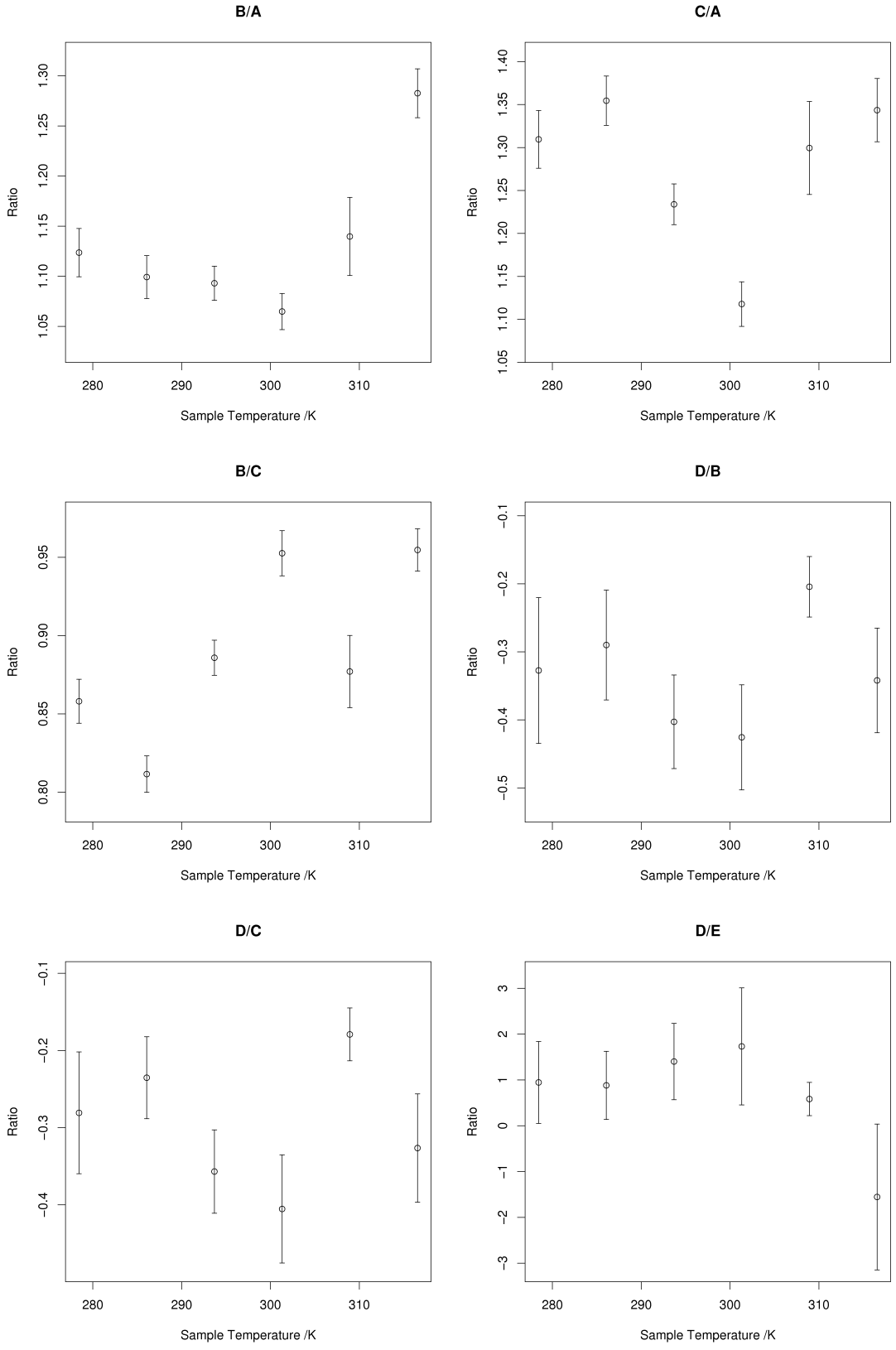


FIGURE 10.8: The plot of ratios between fitted parameters for the 7.5 mM B15C5 temperature dependence dataset

10.3.2 Salt effect on surface orientation

To extend the investigation carried out on the NB15C5 samples, the salt dependence experiments were completed for B15C5. It is hoped the results will show little effect in surface orientation with the added salts. This will confirm the salt effects seen in the past[104] were due to impurities rather than the crown compounds.

Table 10.8 shows the ratio between fitted values for the salt data. These results have also been plotted, shown in Fig. 10.9.

Molar equiv of salt	B/A	C/A	B/C	D/B	D/C	D/E
1x LiCl	1.117±0.103	1.237±0.150	0.903±0.067	0.324±0.380	0.293±0.310	1.288±6.001
10x LiCl	1.108±0.123	1.234±0.177	0.898±0.081	0.302±0.546	0.271±0.440	1.243±9.255
1x NaCl	1.153±0.073	1.249±0.105	0.923±0.047	0.378±0.114	0.349±0.097	1.357±1.467
10x NaCl	1.154±0.035	1.204±0.053	0.958±0.024	0.341±0.075	0.327±0.069	1.190±0.914
1x KCl	1.116±0.029	1.387±0.038	0.804±0.015	0.306±0.114	0.246±0.074	0.847±0.874
10x KCl	1.319±0.059	1.534±0.079	0.860±0.025	0.366±0.071	0.314±0.052	1.317±0.917
None	1.093±0.017	1.234±0.024	0.886±0.011	0.403±0.069	0.357±0.054	1.402±0.834

TABLE 10.8: Non-linear fitting values determined from the 7.5 mM B15C5 salt dependence dataset after applying a $+3^o$ offset to the output polarisation

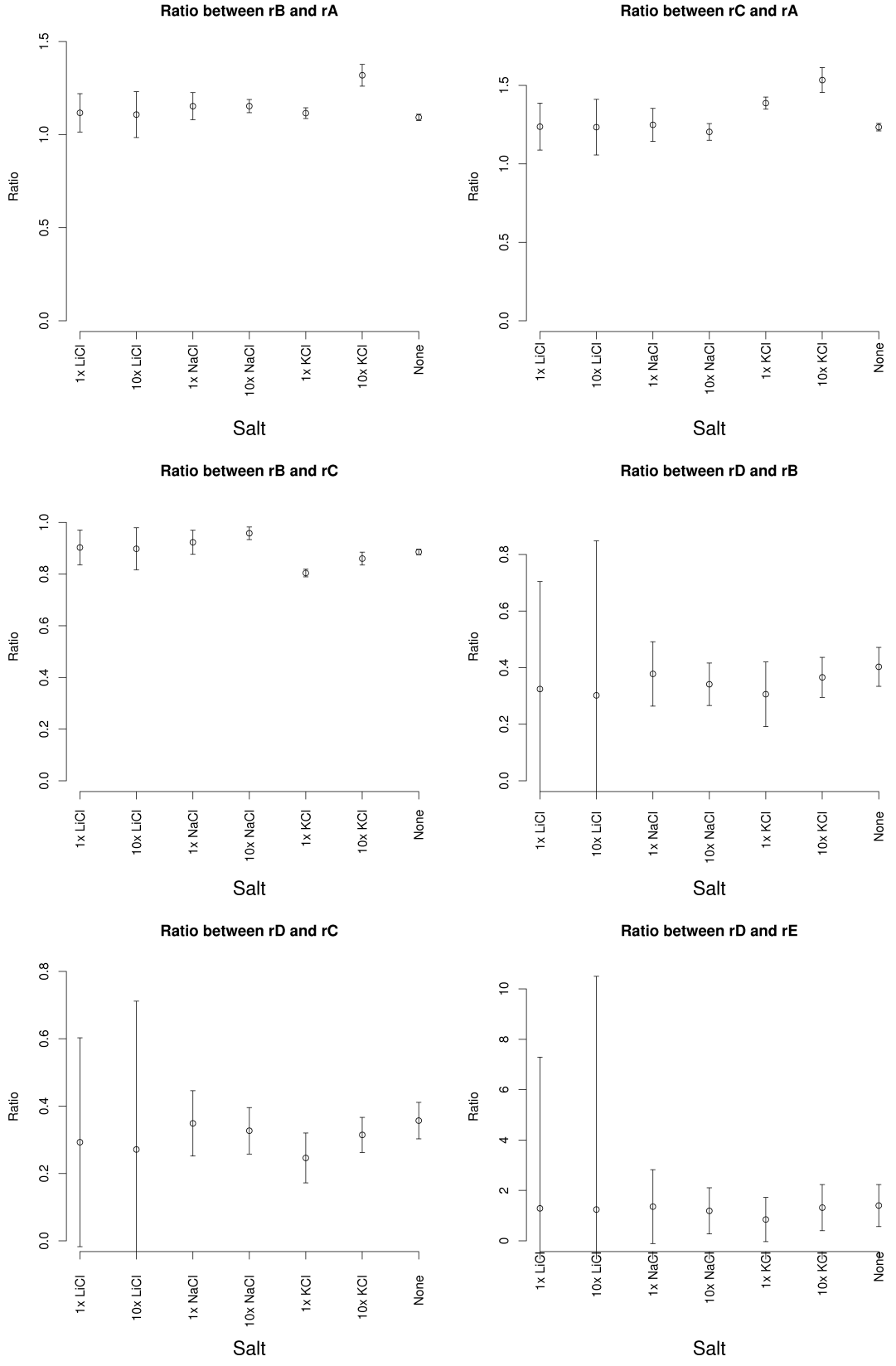


FIGURE 10.9: The plot of ratios between fitted values for the 7.5 mM B15C5 salt dependence datasets

As with NB15C5, these results show very little change in surface orientation as the sodium and potassium salt is added. This furthers the argument that the salt effects seen in the previous work were caused by impurities in the sample rather than from the compound itself. This impurity may be present in the brown oil observed during the purification process, unfortunately this was not collected and therefore could not be analysed further.

The plots also show that D was poorly determined with the addition of the lithium salt. These results suggest that χ_{xyz}, relating to D and E, is not significant in this fitting. This was investigated via an ANOVA study on the '1x equiv LiCl' data, the results of this are shown in table 10.9.

		ANOVA	
Model	RSS	F test	Pr(>F)
A real DE not included	0.0041		
ADE real	0.0023	68.415	7.09e-16
AD real / E complex	0.0021	7.979	0.006
AE real / D complex	0.0023	0.451	0.502
A real / DE complex	0.0019	23.225	1.175e-05

TABLE 10.9: The ANOVA results from the 1x equiv LiCl salt addition to the B15C5 data, determining the significance of the χ_{xyz} term

The ANOVA results show the χ_{xyz} term, and therefore D and E, are significant and therefore should still be included in the fitting. The ANOVA study suggests that the optimal fitting model would be for D and E real, whereas with the previous results these have been considered complex. This may be due to the lithium ion binding with the crown and therefore disrupting the surface although when running these samples the signal levels were very low. Due to this low signal the rise in intensity at P/S, corresponding to the χ_{xzy} term and therefore D, is also low and therefore may be less well determined.

10.3.3 Adsorption isotherms

As shown in the adsorption isotherm results, the B15C5 compound appears to be 'locking' into the surface rather than undergoing an equilibrium with compounds in the bulk, as seen with the other substances investigated in this project. The observed signal remains constant with dilutions until the concentration has dropped to approximately 1 nM. This was investigated further with a number of other experimental methods.

10.3.3.1 Imagine ellipsometry

To investigate the air/liquid interface further a 7.5 mM B15C5 solution was studied via imaging ellipsometry. The imaging ellipsometer uses a laser source to illuminate an area of the sample with a single wavelength. The beam is transmitted to a CCD camera where an image of the surface is produced. If sections within the illuminated area have different optical properties the detected beam will contain modified properties, which are shown in the camera image. One of the difficulties with using the imaging ellipsometer is that only one cross-section of the image can be in focus at any one time. The ellipsometer software attempts to overcome this by taking a number of images over a range of focal points, averaging on that set by the user. This combination of pictures are then used to generate the image with a wider range of focus. Another problem with using the imaging ellipsometer to observed properties at the air/liquid interface is that any artefacts present at the interface are able to move across the liquid surface. If the artefacts are moving faster than the capture rate of the camera or if at such a lower concentration that they rarely pass by the camera then they will not be observed.

The 7.5 mM B15C5 was loaded into ellipsometer and left to settle for an hour. Once settled, the parameters of the ellipsometer were configured to give a high contract image of the B15C5 surface and an image recorded, this is shown in Fig. 10.10. The sample was then left overnight and another image taken, shown in Fig. 10.11. It was hoped this would show if the surface was evolving with time.

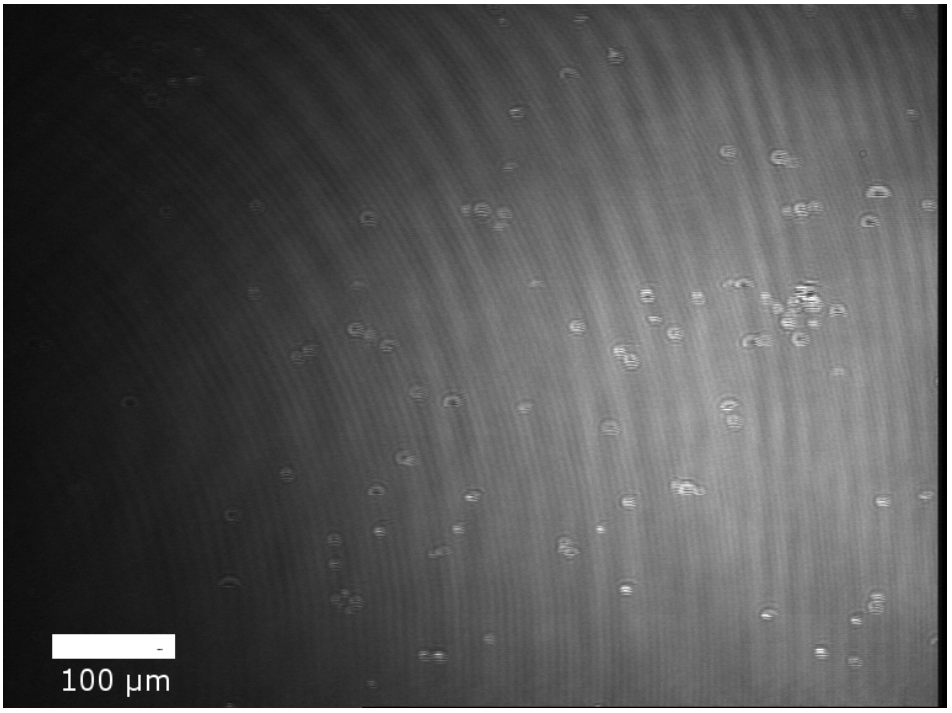


FIGURE 10.10: An image captured from the imaging ellipsometer of the B15C5 surface after one hour

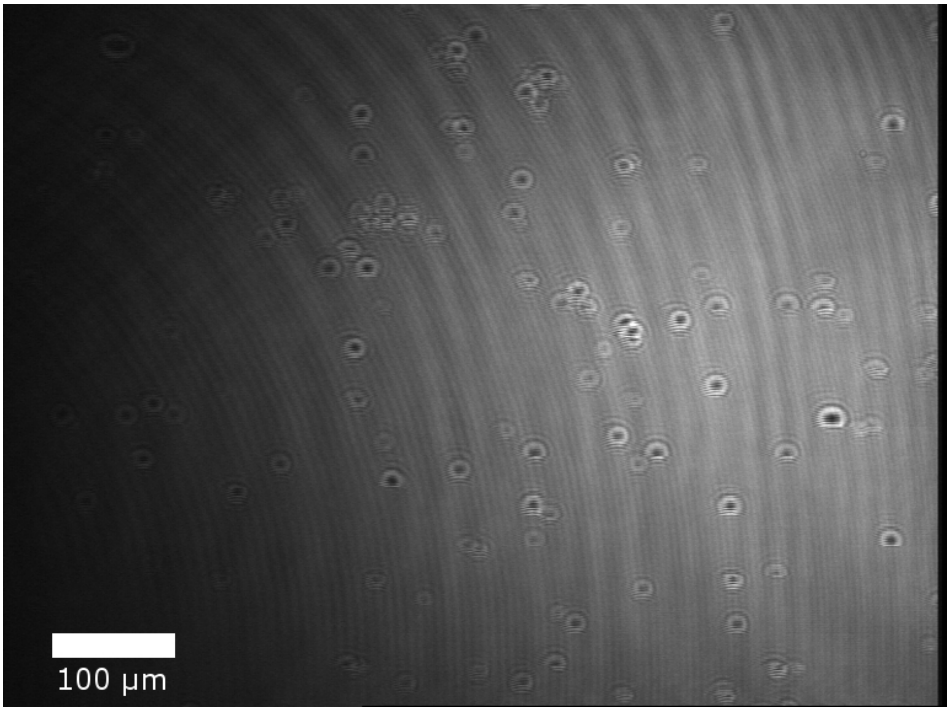


FIGURE 10.11: An image captured from the imaging ellipsometer of the B15C5 surface after approximately fourteen hours

The two images show the presence of clusters or islands on the liquid surface. These islands, not present on a sample of pure water, are measured to have a diameter of 7 μm. The formation of these islands may be a factor in explaining the constant signal in the adsorption isotherms as the concentration is lowered. Should it be more favourable for the molecule stay associated with the island, the molecules will not de-adsorb into the bulk as the concentration is lowered. This will give a constant surface concentration equal to that of the maximum surface concentration obtained from the given sample. This has been tested further with a subsidiary SHG experiment aimed to investigate the signal intensity as a factor of maximum sample concentration.

A second adsorption isotherm was carried out on a 0.5 mM B15C5 solution at 20^oC. The solution underwent a number of 5 ml serial dilutions with pure water until the signal dropped to a level expected for pure water. These results are plotted in Fig. 10.12.

This plot shows that the less concentrated solution has a lower observed signal strength which remains constant with dilutions, as with the previous results. The signal drops at the same concentration of 1 nM for both adsorption isotherms. As the second harmonic signal strength is proportional to the coverage, shown in Eq. 6.41, these results suggest the surface coverage is independent of bulk concentration above 1 nM. The different between these two observed intensities indicate the surface coverage of the B15C5 solution is determined by the highest bulk concentration the given solution has occupied.

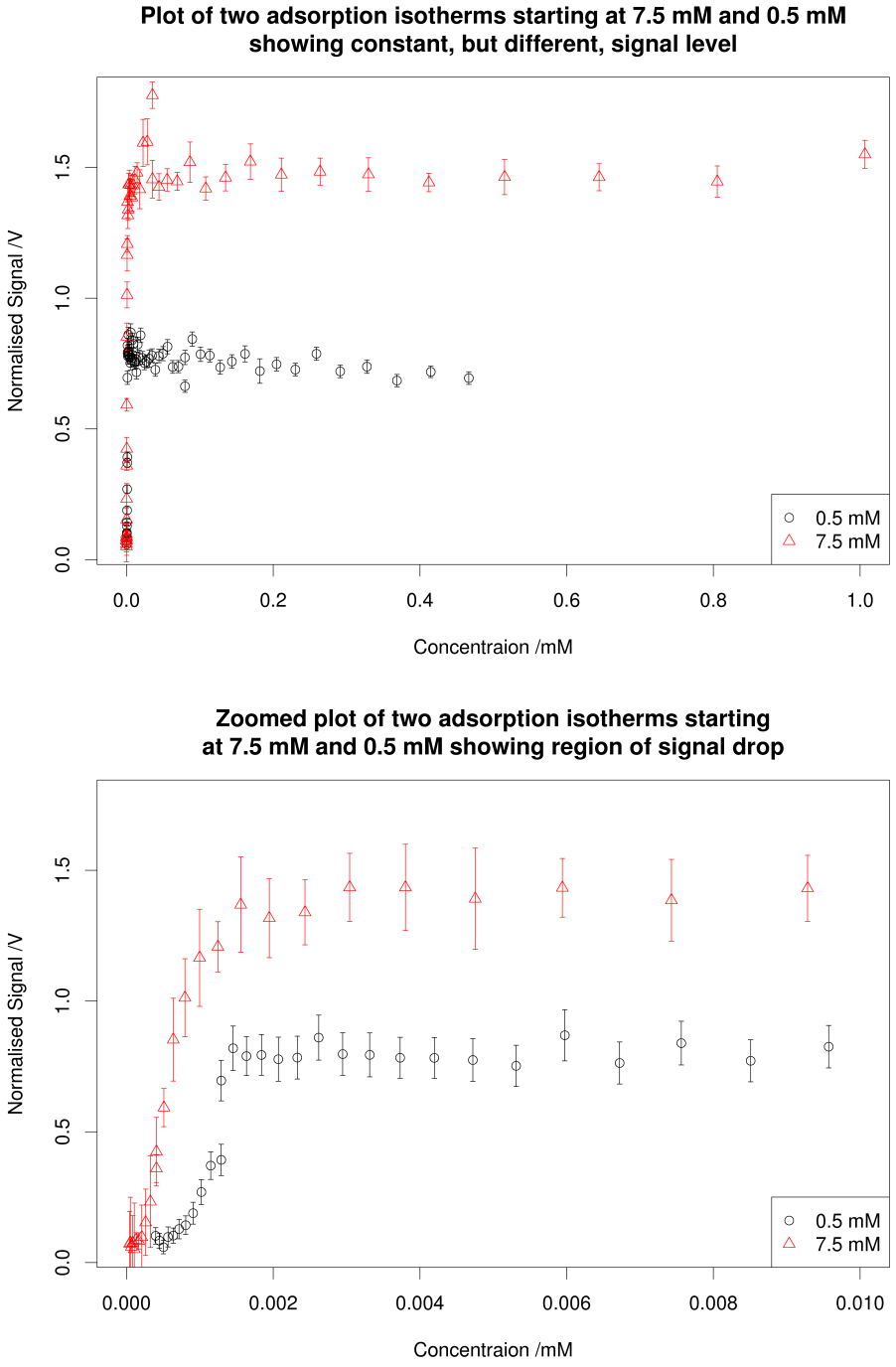


FIGURE 10.12: A plot of the two adsorption isotherms completed at 20^{o}C with starting concentrations of 0.5 mM and 7.5 mM. Top: Concentration range of 0-1 mM. Bottom: Concentration range of 0 - 0.01 mM

10.3.3.2 Surface tension

Previous work has been carried out to study the surface tension of B15C5[132, 133]. After investigation, discussed below, it is thought these results do not account for all processes occurring at the B15C5 surface.

The initial surface tension study looked at the formation of the islands described above. Previous measurements had been recorded as soon as the sample had been prepared, and had not shown changes over time. To record these changes over time a 15 mM solution of B15C5 was prepared, before loading into the measurement dish the solution was shaken vigorously to disrupt any islands formed within the volumetric flask. An initial average surface tension measurement was recorded via the torsion balance. The solution was allowed to settle for ten minutes before another average measurement was obtained, this process was repeated until the measurements had not changed with time. Table 10.10 and Fig. 10.13 shows the results of this experiment. These results show there is a significant change in surface tension over time, accounting for the formation of the islands seen in the imaging ellipsometry results. The settling time has been measured as 40 minutes.

Settling time /minutes	Surface tension /mNm^{-1}
0	56.6
10	54.7
20	53.6
30	53.0
40	52.9
50	52.9

TABLE 10.10: Surface tension measurements over time for a 15 mM B15C5 solution at 20^{o}C

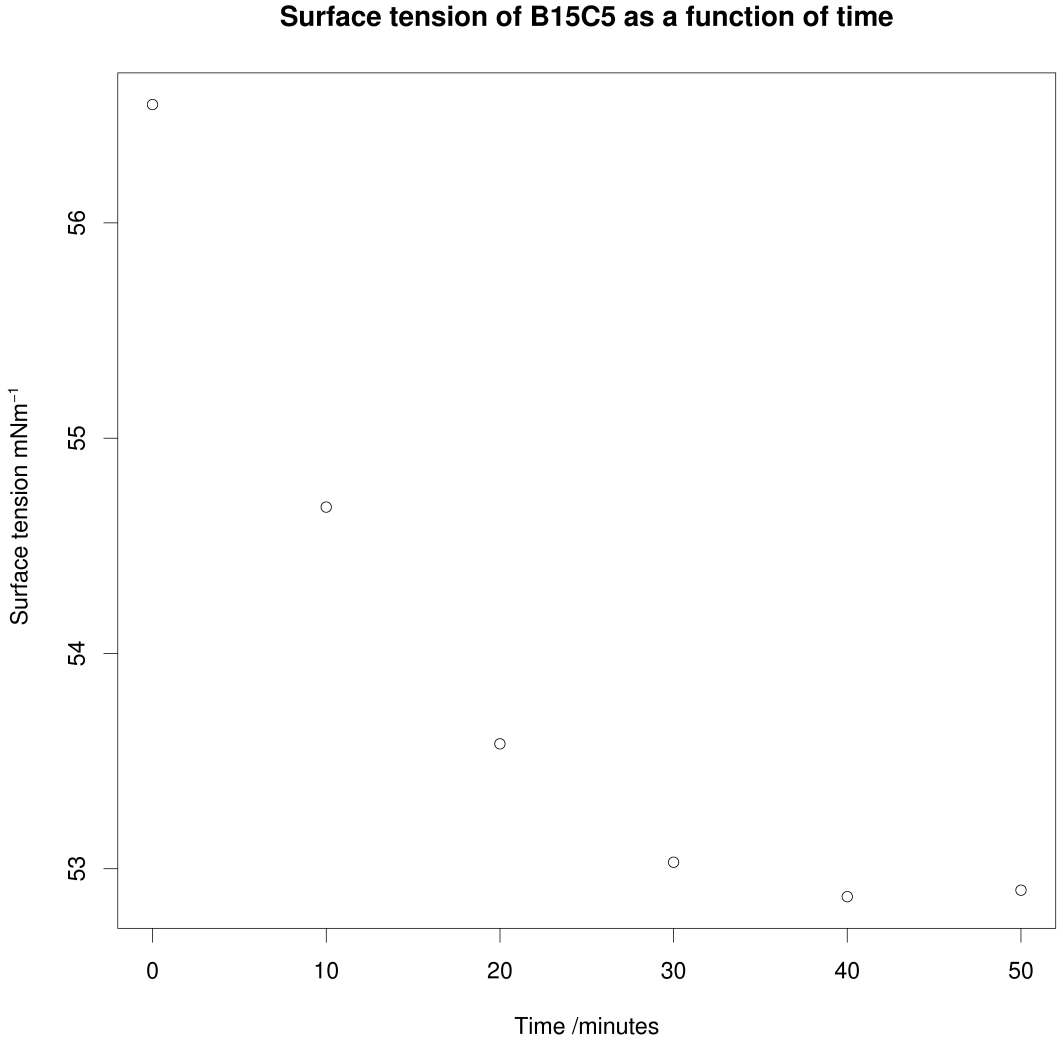


FIGURE 10.13: Plot of surface tension measurements made over time for a 15 mM B15C5 solution at 20^{o}C

Once an understanding of the settling time has been obtained, the initial and settled surface tension was measured as a factor of bulk B15C5 concentration. These experiments were carried out by first preparing a B15C5 solution and loading it into the measuring vessel. An initial average surface tension was recorded, made from at least three measurements. The sample was covered and left to settle for at least an hour before an average settled surface tension was recorded (again from at least three measurements). The sample was then diluted and the measurements repeated. Due to the time-scales of the experiment (at least two hours to prepare the sample and one and a half hours between dilutions including measurements) a maximum of four concentrations could be collected over a single day. As it was thought the sample would decompose over night the samples were not retained. Therefore this experiment was complete in three batches: 10 - 5 mM, 5 - 1 mM and 1 - 0.3 mM. The results from these experiments are shown in table 10.11.

		Surface tension /mNm^{-1}	
Exp. No.	Concentration /mM	Initial	Settled
1	10.03	59.1	49.25
	7.29	52.7	49.0
	4.78	54.55	51.25
2	4.63	64.38	55.9
	2.73	58.5	54.0
	1.45	58.1	54.25
	0.95	60.0	57.5
3	1.06	70.88	65.45
	0.53	67.0	65.25
	0.28	68.88	64.75

TABLE 10.11: Initial and settled surface tension measurements from three B15C5 solutions undergoing serial dilutions at 20°C

As discussed in the imaging ellipsometry section, it has been shown the surface coverage is related to the highest concentration a given solution has acquired. This can be verified via the surface tension results. If this theory is correct only the first 'initial surface tension' measurement from each experiment should align with that of a fresh solution of that concentration (as the subsequent concentrations have been produced through dilution). An experiment to validate this was conducted. A number of independent 'fresh' B15C5 solutions were prepared with concentrations 10.0, 7.5, 5.6, 2.8 and 1.0 mM. Each of these were shaken vigorously before the initial average surface tension measurements were recorded. The results from these experiments are shown in table. 10.12.

Concentration /mM	Surface tension /mNm^{-1}
10.0	58.6
7.5	60.2
5.6	62.0
2.8	64.3
1.0	70.4
0 (pure water)	72.8

TABLE 10.12: Surface tension measurements from independent B15C5 solutions at 20^{o}C

By obtaining the surface tension at various concentrations the surface excess can be calculated using Eq. 10.2[134].

$$d\gamma = -RT\Gamma_s \frac{dc}{c} \tag{10.2}$$

where γ is the surface tension, Γ_s is the surface excess of the species, R is the gas constant, T is the temperature and c is the concentration of the bulk. From this equation it can be seen there is a linear relationship between surface tension, γ, and *ln(c)*.

The results have from table 10.12 and 10.11 have been plotted in this form, shown in Fig. 10.14.

From this plot it can be seen that the surface tension of the fresh solutions matches closely to that of the first solution of each experiment where the solution was allowed to settle. The deviation is due to the 'fresh' solutions being allowed to settle for a short period (less than 5 minutes), as these results were collected before the settling period was fully understood. It can also be seen that all the series have very similar gradients which, as shown in Eq. 10.2, shows the surface excess and therefore area per molecule is effectively constant between series. Between the initial reading of the fresh and settled solutions it can be seen there is an approximate shift of 6 mM in the surface concentration. A further shift of approximately 2 mM is seen between the initial and settled readings of these solutions. This shows that although the number of molecules at the surface is increasing the area they are occupy is remaining constant. This supports the evidence of the formation of islands, allowing molecules to move into an island rather than sitting freely on the surface

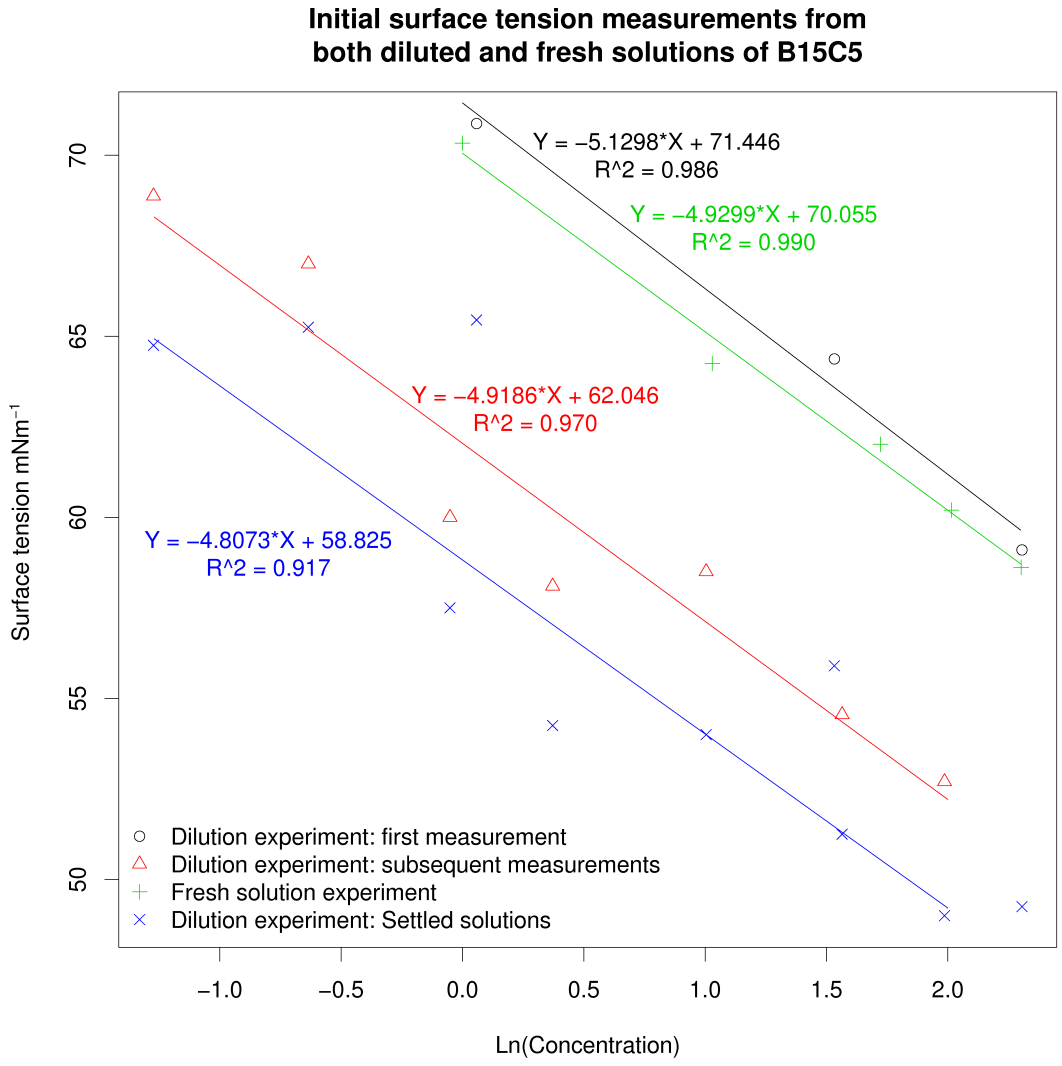


FIGURE 10.14: Plot of Ln(concentration) against surface tension for both dilution and fresh B15C5 solutions measured at 20^{o}C

10.3.4 Chirality

When fitting the polarisation scan experiments to Eq. 6.21 a poor fit was seen. The fitting equation was modified to include the χ_{xyz} term, which relates to chirality in the system, and a significant improvement to the fitting was observed. As B15C5 is achiral, the chirality must related to another aspect of the system. Flack *et al.*[135] discuss the introduction of chirality in crystal structures from achiral molecules. It is through the islands formed on the liquid surface are formed in this way, creating chirality in the system. It is excepted that these crystals would form a racemic conglomerate, a mix of both enantiomers. If the laser irradiates a large number of islands per shot then the chirality would not be observed. To determine the number of islands detected within each laser shot an estimate of the size of the laser beam at the surface is required. This was calculated using heat sensitive paper and irradiating it with a few laser shots. The size of the burn was measure under a microscope and referenced to a reticle. The diameter of the beam was measured as 0.2mm, although assuming a Gaussian distribution of power this may be an over estimate. If it is assumed the laser is perfectly circular this gives a cross-sectional area of 31.4 nm^2. Using the images collected from the imaging ellipsometer an estimate of islands within this area can be determined. This was calculated as approximately 7 islands per shot. As the number of islands observed from a single pulse is low, it is unlikely this small section of the surface will be a racemic conglomerate. Should the chirality seen in these experiments be due to these islands then this may explain a signal containing the chiral component.

Papakostas *et al.*[136] also discuss planar 'structural' chirality from achiral compounds. The chirality here is due to the asymmetrical ordering at the surface, such that if the components are randomly distributed the chirality is not seen. The adsorption isotherm results have already shown that B15C5 is locking into the surface; this locking may also relate to ordering at the surface. Should this be the case, this form of planar 'structural' chirality may be the source of the significance of the χ_{xyz} component in the fitted results.

10.3.5 Further work

The salt dependence experiments showed no changes in surface orientation with the addition of the salts, although effects had been seen in previous work. It has been suggested the effects seen in the previous work were due to impurities in the sample, filtered out during the extended purification process carried out in this project. During this purification a brown oil was discarded. Further analysis on this oil could be used to identify the compound present in the oil and further experiments could be carried out on these. This may show the effects observed in the previous work and identify the compounds causing these effects. These impurities may be open forms of the crown ring produced during the synthesis of the compound; examples of these are shown in Fig. 10.15. Investigation

has begun into 1-methoxy-2-2-[2-(2-methoxyet hoxy)ethoxy]ethoxybenzene (compound *a* in the figure). The initial results collected during this study has shown no surface activity with the addition of the cations.

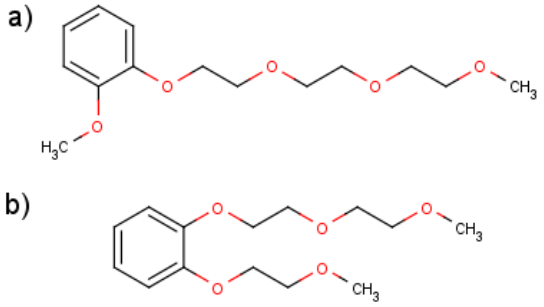


FIGURE 10.15: Potential impurities removed from the B15C5 sample during purification

The ellipsometry results have shown the formation of islands on the B15C5 surface. A full investigation could be completed to get a better understanding of the structure and properties of these islands. Techniques such as Brewster angle microscopy (BAM) can be used in conjunction with compression/expansion isotherms on a Langmuir-Blodgett trough to observe the growth and dispersion of these islands, as carried out with other compounds[137, 138]. The time-scale of the formation of the islands can be studied further by observing the P/S polarisation combination of SHG, corresponding to the χ_{xyz} term, over the 40 minute formation period.

Chapter 11

Design of experiment

11.1 Introduction

The aim of any investigation is to determine how a response is affected by any given factor, and if possible to determine the relationship between them. This is carried out via experiments, observing and detecting the output while modifying one or more of the factors. The aim of DOE is to make these experiments as efficient as possible through use of statistical modelling[139].

A typical investigation will follow a standard life cycle:

1. A hypothesis is declared based on previous knowledge
2. Experiments are completed to generate data to help support the hypothesis
3. The data is analysed and conclusions are made

Should the conclusion deduced from the data not support the original hypothesis, this is modified and the cycle repeated. This is shown in Fig. 11.1.

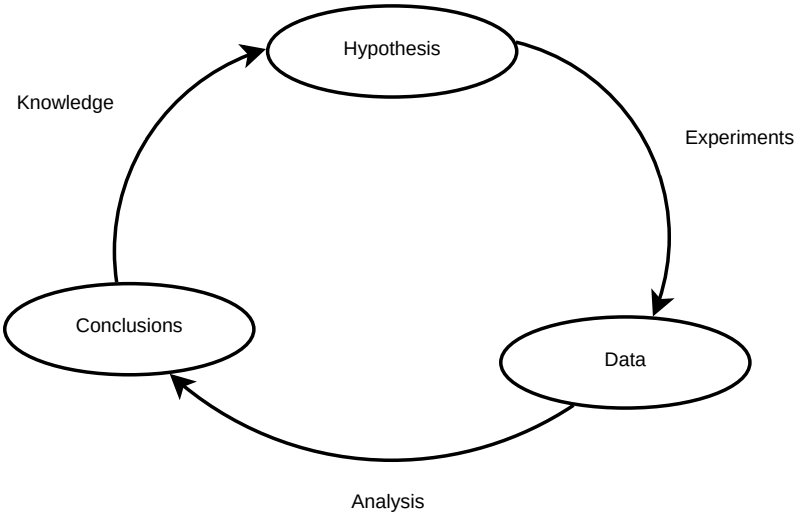


FIGURE 11.1: The experimental life cycle of a typical investigation

The measurements collected during these experiments will contain variation that can cloud the conclusions deduced by the experimenter. DOE adds an additional step into this research cycle, which through modelling, will optimise the experiments to be completed while eliminating the variation. These models attempt to explore the regions of the experimental design space where little information is known. In providing a better understanding of the experimental space, more accurate conclusions can be made and a better understanding is obtained. This is shown in Fig. 11.2.

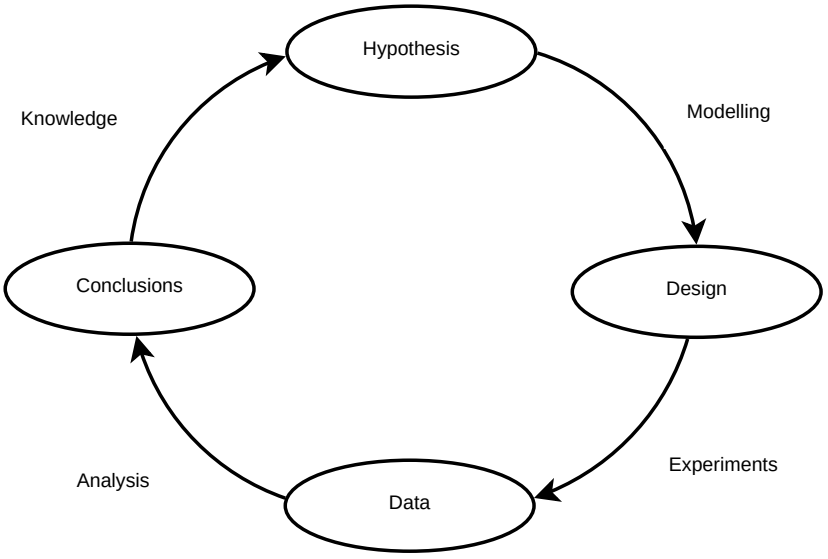


FIGURE 11.2: The experimental life cycle of an investigation including design of experiment modelling

Wu *et al.*[140] give a seven step process in the design and implementation of the DOE methodology, which has been modified and implemented in several case studies[141, 142], and are outlined below:

1. Define objective – Clearly defined the objective of the investigation, such as the scale-up of an experimental procedure in the chemistry domain.

2. Select response – Determine the response or output that is observed and optimised during the investigation. This response should be a measure of the processes occurring and increase the understanding of the system. In the case above, this may be the yield of the experiment.

3. Select factors and levels – Declare which factors, or variables, will be included in the study. It is important to select the significant factors in the process, allowing for maximum knowledge of the system. The values/settings selected for each factor is referred as a level. A combination of factors at a given level is known as a treatment. In the case above, this may be concentrations of starting materials, temperature of reaction and volume of catalyst.

4. Select experimental plan – Determined which design type is suitable for the current point in the investigation life cycle. This can be based on goals of the investigation, prior knowledge of the system and available resources. Some options are pilot studies, factional factorial designs or central composite designs.

5. Complete the experiments – Complete the model produced from the select design which includes a number of treatments. The experiments must be completed with the factors at these given levels and the response recorded. If possible, the experiments should be complete in a random order. This should reveal any systematic errors in the system.

6. Analyse the data – Perform the analysis on the collected data. Typically this will involve fitting the entire data set and determining which factors are not significant with tools such as ANOVA analysis. From this analysis a final model can be presented with response contours based on the selected factors.

7. Make conclusions – Should the investigation have been successful, a model can be presented with optimal parameters for the given process and the most important factors can be highlighted. Should the goal not have been reached new factors and levels for existing factors can be suggested for further investigation.

As the A, B and C parameters fitted to the SHG experiments are non-linear with second harmonic intensity an estimate of these values and the variance associated with them must be understood before a model can be generated. Biedermann *et al.*[143] describe the processes employed in generating the SHG DOE models in detail.

11.2 Second Harmonic Generation models

In the closing months of this project, the Statistics group at the University of Southampton supplied two sets of experimental models for use with the SHG experiment. Each model contained a number of runs or treatments which described the input polarisation, output polarisation and weighting for each run. The weightings were converted to number of laser shots by applying a minimum laser shot count of 250 pulses for a given run and scaling the others to this.

The first set of models were highly reduced, consisting of only six runs for each experiment. This leaving only one degree of freedom in results fitting. These models are shown in table 11.1.

The second set of models contained more runs in each experiment and in theory should provide better estimate of the fitted parameters. The disadvantage to these experiments is that they will take longer to run, although still quicker than the previous methods where no model was used. The second set of models are shown in table 11.2.

Model number	γ /radians	Γ /radians	Weighting	No. of laser shots
1	pi/2	0	0.16	667
	0.44	0.57	0.06	250
	0	0	0.16	667
	2.76	0.79	0.20	833
	2.32	1.33	0.16	667
	2.03	0.79	0.25	1041
2	2.34	1.48	0.16	1000
	0	0	0.15	938
	0.54	0.78	0.04	250
	pi/2	0	0.20	1250
	1.97	0.75	0.25	1563
	2.69	0.88	0.21	1313
3	0	0	0.15	536
	pi/2	0	0.15	536
	2.08	0.99	0.24	857
	2.34	1.39	0.19	679
	0.55	0.72	0.07	250
	2.72	1.02	0.20	714
4	2.59	0.91	0.07	250
	pi/2	0	0.21	750
	0	0	0.16	571
	2.38	1.24	0.14	500
	0.55	0.79	0.17	607
	1.99	0.62	0.24	857

TABLE 11.1: The first, reduced run number, set of optimised SHG experimental models

Model number	γ /radians	Γ /radians	Weighting	No. of laser shots
1	2.22	0.87	0.08	6667
	0.82	1.31	0.09	7500
	2.62	0.04	0.05	4167
	0.33	0.54	0.12	10000
	2.37	1.40	0.003	250
	pi/2	0	0.14	11667
	0.79	1.48	0.04	3333
	3.11	0.02	0.17	14167
	0.27	0.48	0.10	8333
	0.25	0.39	0.05	4167

	2.06	0.71	0.16	13333
2	0.32	0.53	0.06	750
	0.37	0.52	0.05	625
	2.03	0.68	0.11	1375
	2.20	0.87	0.03	375
	0.94	0.91	0.04	500
	pi/2	0	0.13	1625
	0.26	0.50	0.12	1500
	0.79	1.52	0.11	1375
	0.32	0.24	0.02	250
	3.12	0.04	0.17	2125
	2.13	0.79	0.09	1125
	2.61	0.02	0.08	1000
3	2.60	0	0.06	1500
	0.24	0.43	0.07	1750
	2.14	0.83	0.05	1250
	2.36	1.45	0.11	2750
	1.11	0.67	0.01	250
	2.61	0.06	0.06	1500
	pi/2	0	0.12	3000
	0.33	0.54	0.13	3250
	0.99	0.85	0.06	1500
	2.08	0.79	0.05	1250
	1.07	0.77	0.02	500
	3.08	0.07	0.17	4250
	2.03	0.70	0.08	2000
4	2.08	0.88	0.06	5000
	0.34	0.61	0.12	10000
	2.07	0.73	0.07	5833
	0.57	0.01	0.08	6667
	pi/2	0	0.13	10833
	1.04	0.91	0.06	5000
	1.10	0.73	0.08	6667
	2.81	0.58	0.11	9167

0.78	1.55	0.007	583
2.36	1.56	0.10	8333
2.39	1.43	0.003	250
2.71	0.75	0.02	1667
0.02	0.05	0.16	13333

TABLE 11.2: The second, extended, set of optimised SHG experimental models

As seen in the second set of models, when applying the weighting to the number of laser shots some of the values given are experimentally unreasonable. Experimental runs containing more than 2500 are unreasonable to run due to the time required to collect them; this has been seen in the previous studies where evaporation has affected the results. The lowest number of laser shots of 250 was chosen as this included enough data points to obtain a reasonable average, therefore this could not be lowered to accommodate these models. Therefore the models with larger laser shot values than this were disregarded, leaving only model 2 from the second, extended, set.

11.2.1 Model comparison

To compare the results and quality of fit from the two models a number of experiments were carried out. A 75 mM PNP solution was prepared via the standard procedure and a number of SHG experiments carried out using model 1 from the reduced set, referred to as 1.1, and model 2 from the extended, referred to as 2.2. Each model was carried out twice to show any variation between the experiments. During the experiment the water bath was set to 20^oC . Table 11.3 details the experiments in this dataset.

DOE model	Experiment number
1.1	653
1.1	654
2.2	652
2.2	655
Original	317 & 318

TABLE 11.3: The combination of experiments in the DOE model comparison investigation of 75 mM PNP at 20^oC . The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

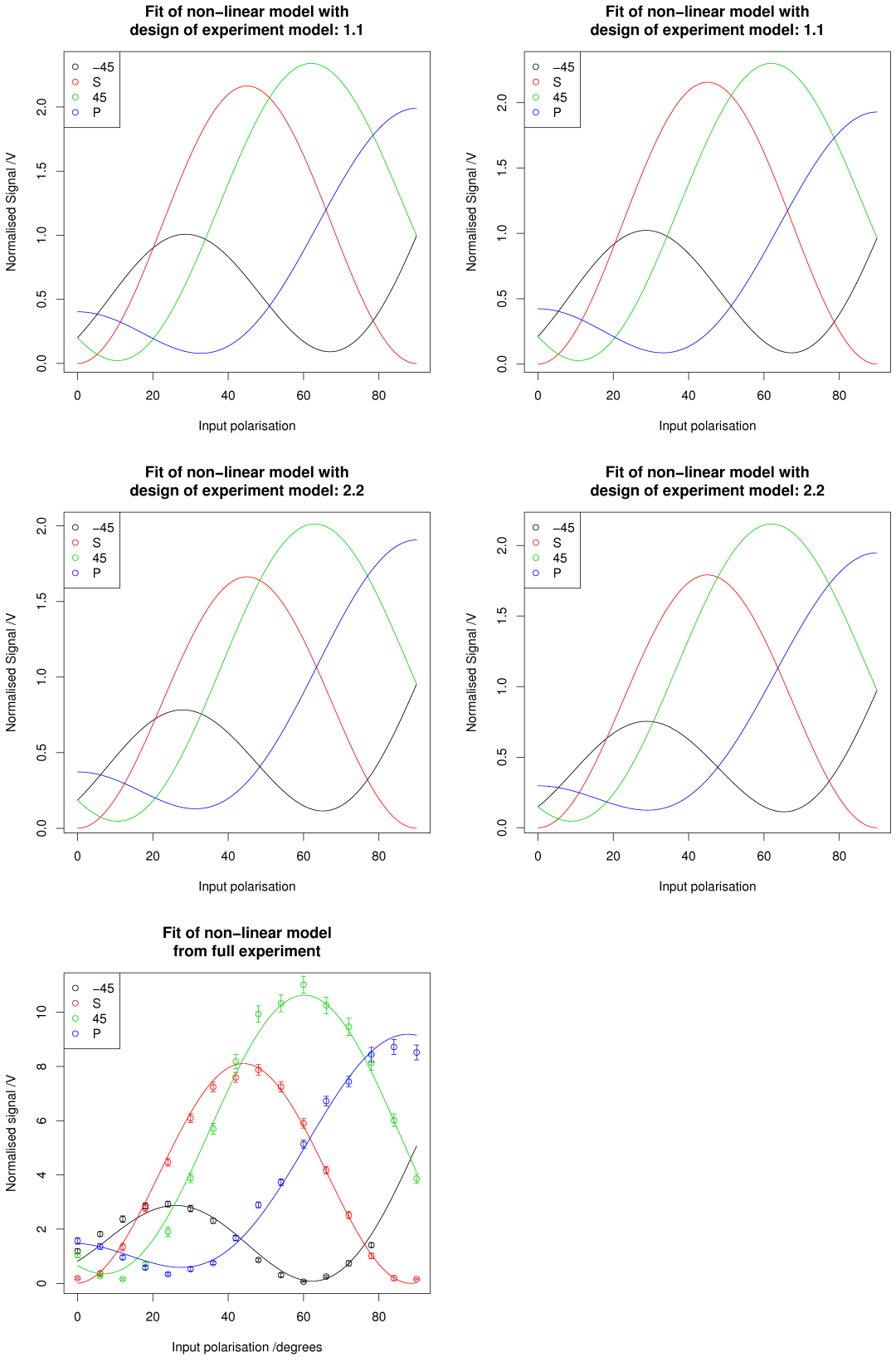


FIGURE 11.3: The plot of fitted curves from the 75 mM PNP at 20^{o}C results from the design of experiment models, compared to the existing experimental data

Table 11.4 shows the ratio between fitted parameters from these experiments and Fig. 11.3 shows the fitted curves, compared to the previous 75 mM PNP data at 20^{o}C . The fitted curves do not show the experimental results as these are not completed at the standard output polarisations. Fig. 11.4 shows a comparison plot of the ratio between fitted values for the models.

Model	Experiment number	B/A	C/A	B/C
0 (Full experiment)	317	2.489±0.013	2.388±0.011	1.042±0.002
1.1	654	2.135±0.102	2.257±0.121	0.946±0.020
1.1	653	2.223±0.101	2.318±0.121	0.959±0.019
2.2	652	2.262±0.060	2.112±0.055	1.071±0.014
2.2	655	2.562±0.116	2.456±0.102	1.043±0.019

TABLE 11.4: The ratio between fitted values from the 75 mM PNP at 20^{o}C from the design of experiment models, compared to the existing experimental data

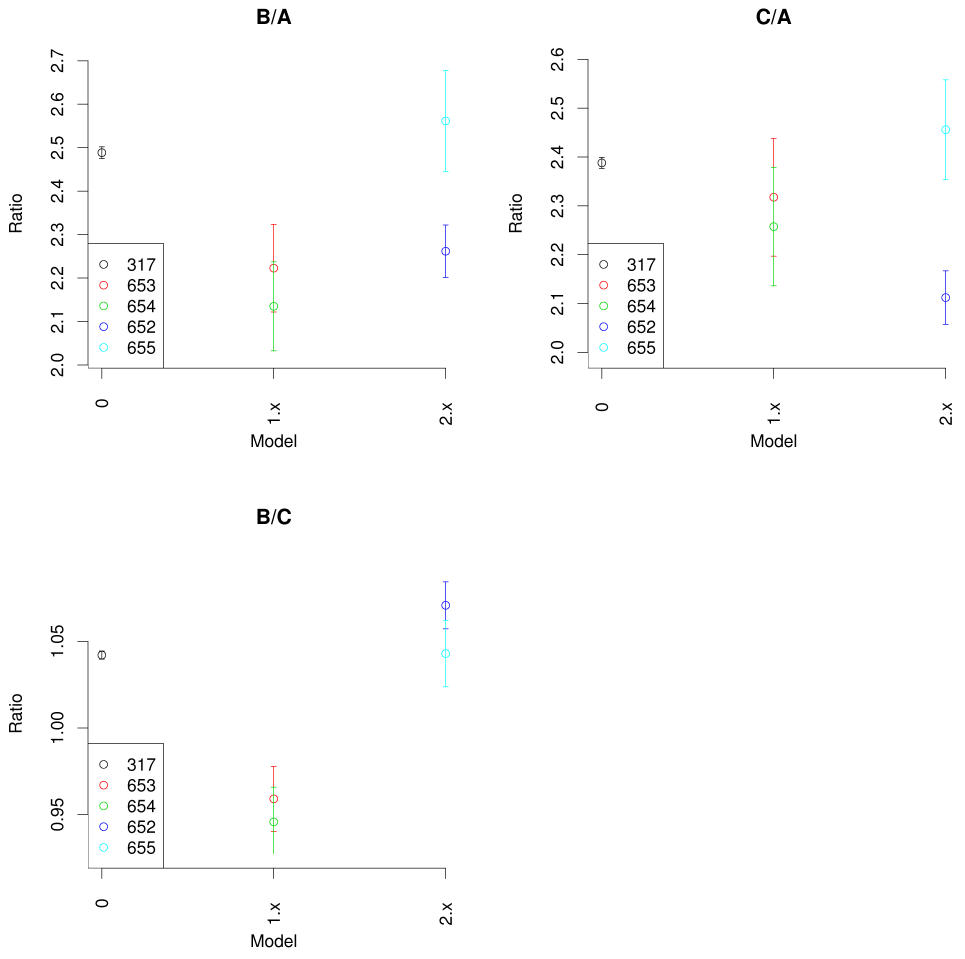


FIGURE 11.4: The plot of ratios between fitted values for the 75 mM PNP at 20^{o}C from the design of experiment models, compared to the existing experimental data (labelled as model=0)

These plots show the values calculated from the two models are very similar to that of the previous experimental results, with the final experiment completed (experiment 654) overlapping well with the original experimental data. As all four experiments were completed with the same sample it was thought there may be some re-ordering of the surface with time which would account for this deviation from the expected results in the earlier experiments, which is discussed below. As shown in the plot, the error associated with the DOE results is greater than that of the original data, which is expected as there is significantly less data available to complete the fit.

The original experiment, labelled 318, took approximately three hours to complete. Using DOE model 2.2 this was reduced to 45 minutes and with model 1.1 the experimental time was just 15 minutes. This shows the significant reduction in time required for the experiments while still capturing the same quality of results.

11.3 Results

The DOE models were used to validate and extend the existing studies carried out on PNP. It is hoped this will highlight the increase in speed and good experimental practise while maintaining or improving the quality of the result.

11.3.1 Surface formation

An experiment to validate the phase transition seen with respect to temperature in PNP, discussed in section 8.4.1.1, was carried out using the DOE models. The experiment used a repeat of the previous procedure, varying the water bath from 0^oC to 50^oC in 10^oC steps. The study consisted of two datasets, the first running from low-to-high and the second from high-to-low. As these experiments could be carried out over a much shorter time scale a single solution was used to complete one temperature range (either low-to-high or high-to-low). A concentration of 75 mM was used throughout these experiments. Table 11.5 describes the dataset of these results.

These experiments were executed and the results fitted to Eq. 6.21 as with the previous experiments. Fig. 11.5 shows a plot of the ratios between fitted values for the two datasets. These results do not show the expected phase transition at 24^oC, therefore another process is occurring at the interface.

Rather than plotting against water bath temperature, the ratios were plotted against relative experiment start time in minutes. This plot is shown in Fig. 11.6. The plot shows there is a time factor related to the surface orientation, which is affected by the temperature. This factor was studied further via a number of additional experiments.

Dataset	Water bath temp. /^{o}C	Experiment number
Low-to-high	0	658
	10	659
	20	660
	30	661
	40	662
	50	663
High-to-low	50	683
	40	684
	30	685
	20	686
	10	687
	0	688

TABLE 11.5: The combination of experiments used to generate the initial temperature dependence of surface orientation dataset for 75 mM PNP using DOE model 2.2. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

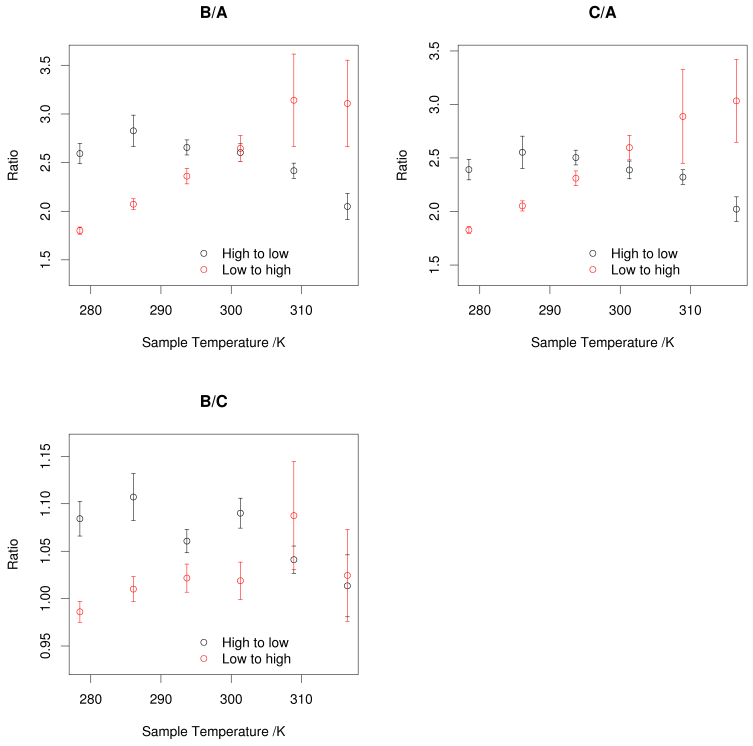


FIGURE 11.5: Plot of ratios between fitted values against water bath temperature for the two 75 mM PNP temperature dependence datasets (high-to-low and low-to-high) using DOE model 2.2

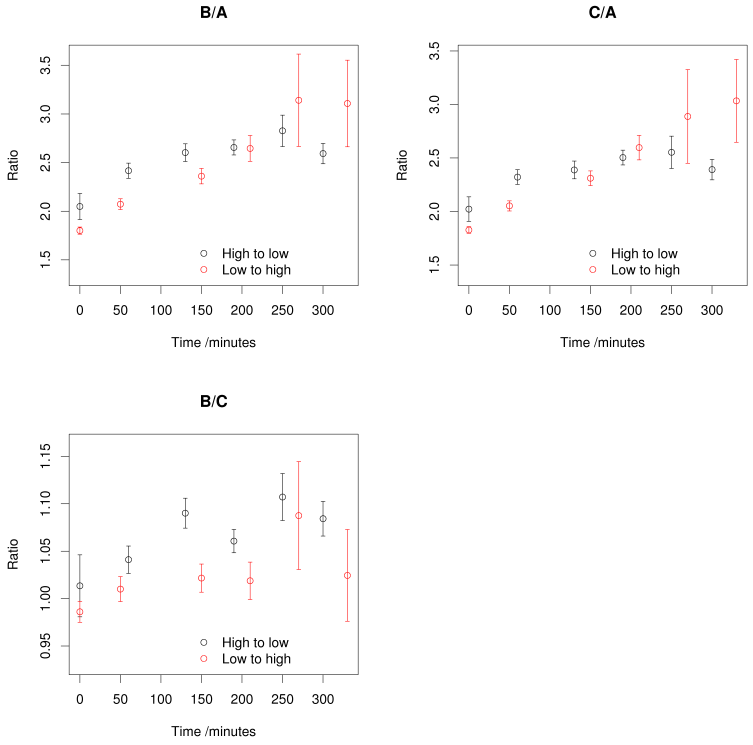


FIGURE 11.6: Plot of ratios between fitted values against relative start time for the over time 75 mM PNP experiment using DOE model 2.2

An experiment was completed to determine the changes in surface orientation with time. This would not have been possible without the DOE models, as it is thought the formation time scale is equivalent to that of a single experiment using the old methods. A 75 mM PNP solution was prepared and loaded into the SHG rig. The software was configured to run the DOE model 2.2 every one and a half hours over a fifteen hour period. Table 11.6 details the experiments in this dataset. The results from these experiments were fitted as with the previous results. Fig. 11.7 shows a plot of the ratio between fitted values against time; the last experimental result has been omitted as the solution had evaporated by this time.

Relative start time /minutes	Experiment number
0	668
90	669
180	670
270	671
360	672
450	673
540	674
630	675
720	676
810	677
900	678
990	679
1080	680
1170	681

TABLE 11.6: The combination of experiments in the surface formation time investigation of 75 mM PNP at 20^{o}C using DOE model 2.2. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

These results show a steady change in the ratios between fitted values, and therefore surface orientation, with respect to time before 300 minutes. After 300 minutes the ratios remain constant within experimental error. These results show there is a time factor to be considered when investigating PNP at the air/liquid interface, relating to the formation of this surface. It would not have been possible to observe this factor without use of the DOE models due to the time scales required to complete an experiment; previously requiring up to three hours to complete.

These results may also explain the deviation of the initial experimental results from the fitted curves seen in the PNP adsorption isotherm study, discussed in section 8.3.3. This factor cannot be investigated further with the PNP adsorption isotherm results as the time between loading the sample and executing the experiments is not known. When running these experiments it was assumed there was no time factor to be consider as it

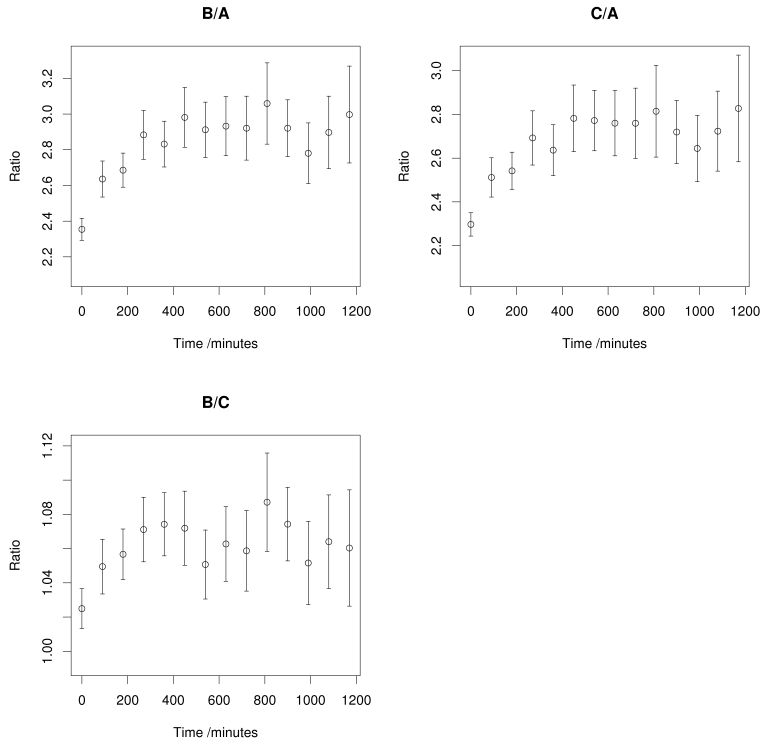


FIGURE 11.7: Plot of ratios between fitted values against relative start time for the two 75 mM PNP temperature dependence datasets (high-to-low and low-to-high) using DOE model 2.2

had not previously been observed, therefore the time between loading and running the sample varied between experiments.

11.3.2 Temperature dependence of PNP surface orientation

As shown above, the initial data, aimed to validate the phase transition of 75 mM PNP observed at a water bath temperature of 24^oC, did not show the expected results. It has been shown this is related to a settling time of the surface. Before running the DOE temperature dependence experiments a second time, a further investigation was completed to determine if the settling time was required between each change in temperature or if it is only observed with a fresh solution. A 75 mM solution was prepared and loaded into the reaction vessel, with the water bath set to 20^oC. The solution underwent a number of experiments using DOE model 2.2 run once every one and a half hours until the surface had settled. The water bath was then heated to 30^oC, the solution again underwent a number of experiments using DOE model 2.2 run once every one and a half hours. Table 11.7 details the dataset of this investigation. These results were fitted using the R script as with previous results. Fig. 11.8 shows the plot of ratios against temperature from the data of this investigation.

Water bath temp. /^{o}C	Relative start time /minutes	Experiment number
20	0	689
20	90	690
20	180	691
20	270	692
30	0	693
30	90	694
30	180	695

TABLE 11.7: The combination of experiments in the surface formation time investigation of 75 mM PNP at multiple temperatures using DOE model 2.2. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

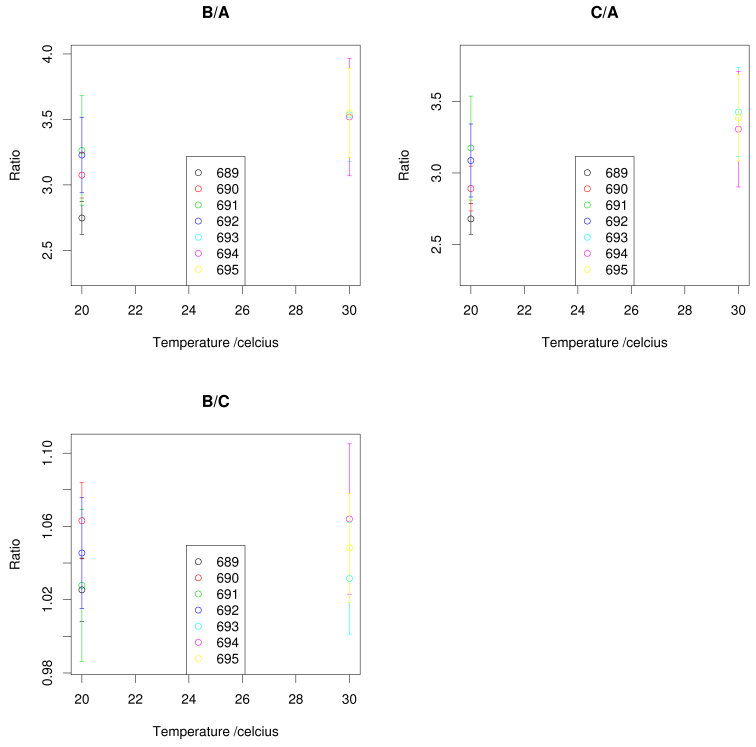


FIGURE 11.8: Plot of ratios between fitted values against temperature for the 75 mM PNP time dependence dataset included a change in temperature, legend shows experiment number

These results again show the settling time of the PNP at the initial temperature. The results show no change, within experimental error, in the second temperature. This shows any further re-orientation, relating to the change in temperature, occurs a time-scale fast than the DOE experiments and therefore will not be considered.

Now a full understanding of changes occurring on the DOE experimental time-scale has been obtained, the validation of the previous temperature dependence results can

be carried out. A 75 mM PNP solution was loaded into the experimental vessel and allowed to settle for five hours. After this time a number of experiments, using DOE model 2.2, at water bath temperatures 0, 10, 20, 30, 40 and 50^{o}C were carried out. Table 11.8 details the experiments in this study. These results were modelled as with the previous results and the fitted parameters determined. Fig. 11.9 shows the plots of fitted ratios for this dataset.

Water bath temp. /^{o}C	Experiment number
0	703
10	704
20	705
30	706
40	707
50	708

TABLE 11.8: The combination of experiments in the study to validate the 75 mM temperature dependence results, completed using DOE model 2.2. The data can be downloaded from http://middleware.chem.soton.ac.uk/shg/experiment/**Experiment number**

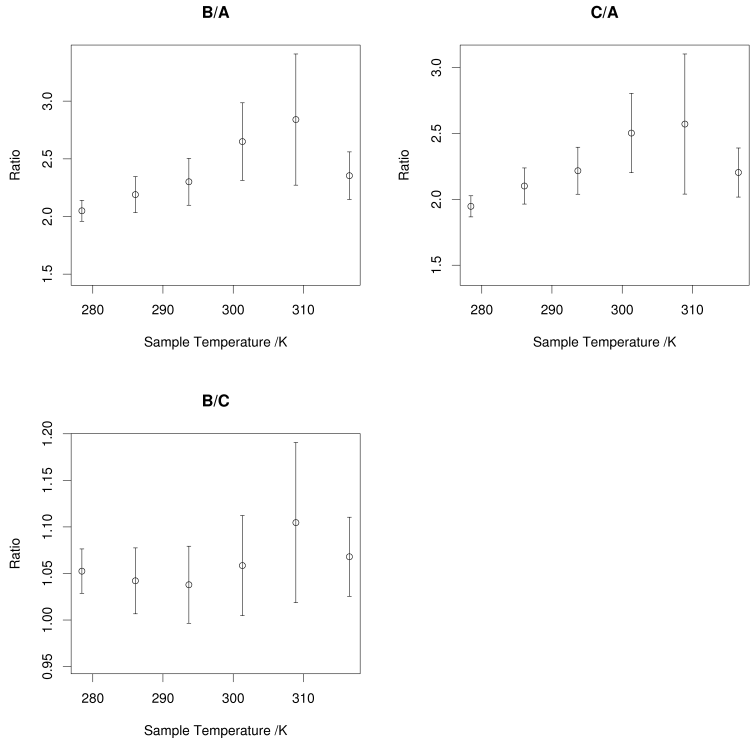


FIGURE 11.9: Plot of ratios between fitted values against temperature for the 75 mM PNP temperature dependence validation dataset collected using DOE model 2.2

If the 50^{o}C data is considered as an outlier, the trends seen in the DOE data matches those seen in the original PNP temperature dependence results, shown in Fig. 11.10.

This may be considered an outlier due to the evaporation issues seen in previous experiments. The absolute values obtained from the DOE dataset are lower than those seen in the original study, which may be a factor of the settling time as the original study did not account for this. The full experiments used in the original study were completed over a similar time-scale to that of the settling process, therefore an average between fresh and settled surface is likely have been collected in these experiments. As in previous experiments the 50°C data is likely to have been effected by evaporation causing the alignment of the laser to change, which is highlighted by the lower signal observed in this experiment. This supports the classification of this data point as an outlier.

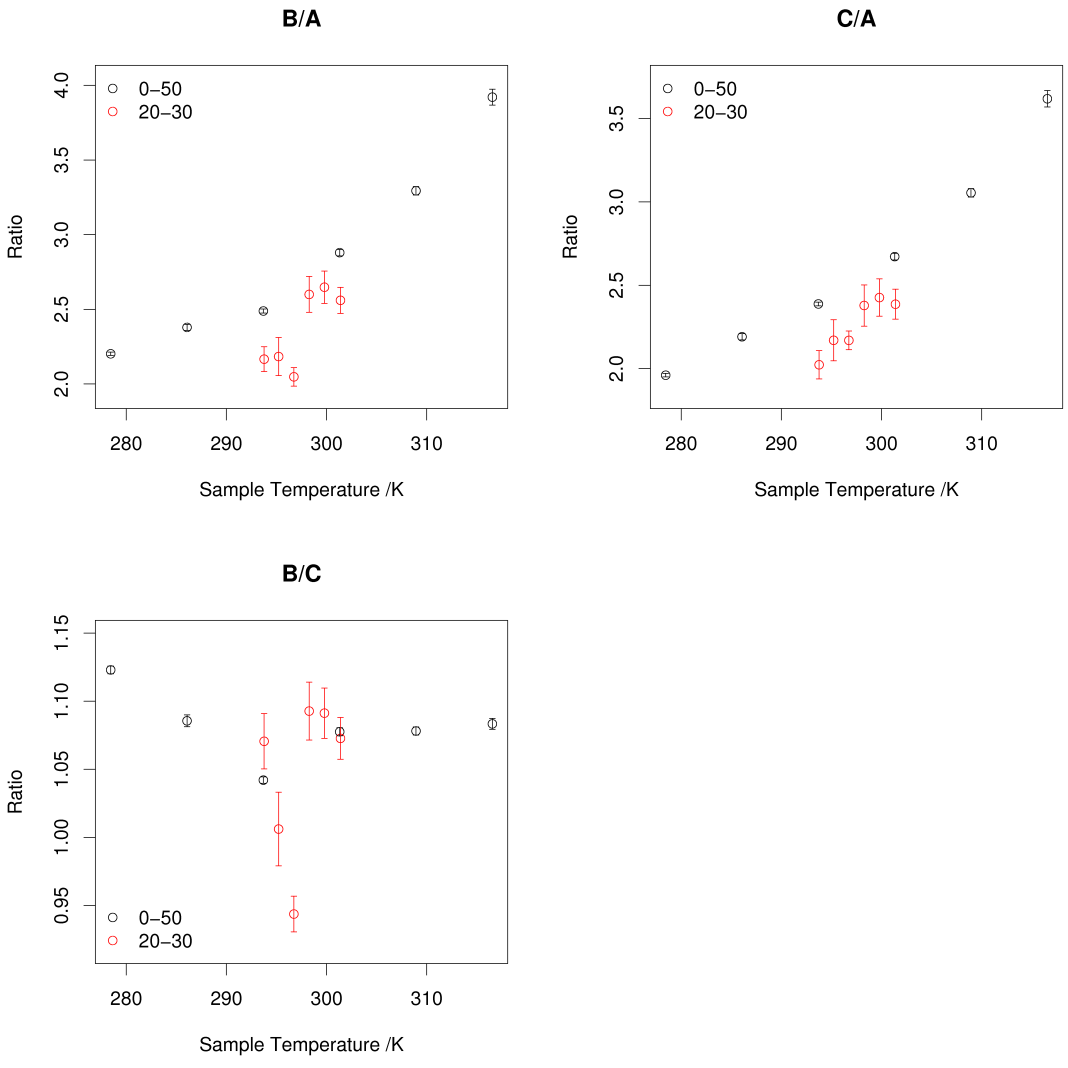


FIGURE 11.10: The plot of ratios between fitted values for the two 75 mM PNP datasets. Black indicates the 0-50 dataset and red the 20-30 dataset

11.4 Conclusions

It has been shown that with the use of the DOE methodology, the SHG experimental process has been greatly improved. The experimental time-scale has been reduced from approximately three hours to fifteen minutes in the most reduced model. In using these new experimental practises a phenomenon has been observed, where previously not possible in this detail. The data collected from the DOE experiments shows a good match to that observed in the previous study. A disadvantage of these models is that there is greater uncertainty of the determined parameters, which is due to the lower degrees of freedom in the fitting process. This trade-off can be investigated further to determine an optimal model within a time-frame while reducing the error to that of an acceptable level.

11.4.1 Further work

As the parameters determined through the SHG experiments are non-linear with response, an initial estimate of these parameters is required before a model can be generated. The DOE models provided by the statistics group were designed specifically for 75 mM PNP at the air/liquid interface at 20^{o}C , although as seen in this investigation these models can be used away from the optimal conditions. To extend this investigation further models could be generated for a number of other compounds and at different concentrations. A comparison of these models can determine how robust any given model is to changing one or more experimental parameter. As seen with B15C5, some compounds will have extra terms required in the fitting and therefore a new group of models would be needed before completing these experiments.

As shown in chapter 4, all aspects of the SHG experiment have been automated and integrated with a work-flow controlling script via the message broker. This can be extended to the generation of the DOE experimental models. The DOE models are procured through an R script, and therefore a Perl wrapper can be used to handle the message broker interaction, as with the A, B and C fitting. The additional step of model generation would then simply be a called via a message to the broker. As an initial estimate of the fitted parameters is required before the model can be generated, a standard experiment would need to be completed and fitted first. Once a model has been determined for a given compound, it can be stored in a database and recalled for further experiments.

Chapter 12

Summary and conclusions

This project has been split into three interlinked sections; the development of a data acquisition system for environmental conditions, applied to the chemistry laboratory, the development of an experimental control system which was has been used in the SHG experiment and the use of this software to complete a surface study of PNP, NB15C5 and B15C5 at the air/liquid interface. In the closing stages of the project, DOE techniques have been used to optimise the SHG experimental practice; this was only possible due to the automation of the experimental apparatus.

An environmental data acquisition system to capture the conditions within the chemistry laboratory has been developed. The system uses a message broker, the IBM Microbroker, to control the flow of messages between components in the acquisition process. The system has been developed in a modular format to allow a plug-and-play like deployment of the system. The core components of the system are:

- Acquisition component – Captures raw data from a given data acquisition card and publishes the raw voltage reading. During the work the standard acquisition card has been the Labjack UE9, although other cards have been included to show the transformability of the solution.
- Calibration component – Subscribes to messages containing raw voltage readings, performs the calibration and publishes the calibrated values as a new topic on the message broker.
- Change component – Monitors the calibrated data produced by a given sensor and filters small changes. This component was introduced after first testing as it had been noted a large amount of repeated data had been collected. The change threshold is set by the user, allowing the resolution of the data to be controlled.

- Storage component – Receives data from a given channel, converts it into a format suitable for the selected repository and archives it. A number of storage components have been developed to allow comparison of a number of repository solutions during the work.

In addition to these core components, a number of optional 'plug-ins' have been developed including an alarm system and an extended real-time monitoring solution. The system has been used to monitor a number of laboratories in the School of Chemistry and School of Physics at the University of Southampton. The data collected from these laboratories has been used as metadata, providing additional trust in experimental results as well as providing some explanation in anomalous results.

During the project a number of repository solutions have been implemented and compared. A MySQL database was deployed, using a schema developed in-house, to store both the sensor data and its associated metadata. This was made available through a web interface. An implementation of the OGC SOS was also deployed, which is a global standard describing an API for managing sensors and their data. This allowed for greater interoperability between system but had a larger overhead in deployment and maintenance. Web base solutions such as PaChube and SensorBase were also investigated; these move the maintenance of the repository away from the user although the policy for storage and archiving cannot be controlled.

A remote experimental control system has also been developed; as with the environmental monitoring system this uses the message broker to mediate the flow of data between control interface and experimental apparatus. The control system uses a standard framework which is extend to communicate with a given piece of laboratory equipment. This has been applied to a SHG experimental set up, allowing control over laser source, shutters, half-wave rotation and the dilution of the liquid sample. In automating the SHG experiments, the results have become more reliable and repeatable while allowing the investigator to leave the apparatus unattended.

To provide access to the SHG experimental results and the associated environmental conditions a 'mash-up' website has been developed. This provides a central, and web addressable, access point to all the SHG experimental data. This website has been developed using open source software, showing how a powerful research tool can be developed using existing and open solutions.

With the control and monitoring systems in place in the SHG laboratory, an investigation using these tools could be undertaken. PNP was chosen as an initial compound to study due to its large signal at the selected wavelength and its interest in pollution study. A temperature and concentration dependence study of surface orientation was completed. The temperature dependence showing a phase transition at approximately 293K while showing a steady, although small, change in orientation with concentration. Through

further analysis it was found that the rotation about the z-axis, ψ, should be fixed to 90^o when fitting the experimental results. A number of adsorption isotherm have been complete at temperature between 0 and 50^oC and fitted to a Langmuir isotherm. These results have been used to estimate the Gibbs free energy of adsorption to the surface, $\Delta_{ads}G$, as -20.5 kJ mol^{-1} at 298K. From this an estimate of the enthalpy, $\Delta_{ads}H$, and entropy, $\Delta_{ads}S$, of adsorption to the surface at 298K as -43.6 $\pm$ 2.64 kJ mol^{-1} and 77.7 $\pm$ 9.11 J mol^{-1} K^{-1} respectively

A study of NB15C5 was also carried out. As with PNP, an investigation of temperature dependence on surface orientation was carried out. These results showed that the orientation remained constant with temperature while the ratio between dominant hyperpolarisability terms changed, suggesting a change in the electrical configuration of the molecule. A study of the effect of cations on the surface was carried out with lithium, sodium and potassium, selected as the ionic radius is similar to that of the crown ring size. These results showed little change with the addition of the cations, contradicting results seen in a previous study. It is thought the difference between results is due to impurities in the NB15C5 solutions in the previous study, which were removed in the purification procedure completed in this work. An adsorption isotherm was also completed on a NB15C5 solution at 298 K. These results show a phase transition at a concentration of 4.4 mM. These results were studied further with full polarisation scans at multiple concentrations, showing significant changes in dominant hyperpolarisability terms with concentration. This phase transition was verified by surface tension measurements.

The final compound studied in this project was B15C5. A study of temperature dependence on orientation was completed. These results showing the fitting model needed modification; accounting for the χ_{xyz} term which suggests chirality. The results from the study showed little change in the ratio between the fitted parameters with changes to temperature. As with NB15C5, a salt dependence study was completed. These experiments showing little changes in fitted parameters with the addition of the salts. An adsorption isotherm was carried out on this compound. The results showing the compound was 'locking' into the surface, giving constant signal with dilution down to 1 nM. This was investigated further. Imaging ellipsometry was used to view the liquid surface. This showed the formation of 7 μm diameter islands on the liquid surface. These islands were studied further via surface tension. This study showed the island were formed over approximately forty minutes. The surface excess of the 'fresh' and settled interface being equivalent but showing a shift in surface concentration of approximately 6 - 8 mM.

In the closing stages of this project a study of PNP was completed using DOE methodology to optimise the SHG experimental practise. Two sets of experimental models were provided by the statistics group in the University of Southampton. These models were first analysed for experimental viability then compared experimentally. These models provided experimental results equal to those from the full experiment within experimental error. A model was chosen and the PNP temperature dependence study repeated.

The results collected from this were unexpected but through further analysis suggested the interface required time to settle; this effect was masked in the previous study due to the long collection time required for the full experiments. An experiment was conducted to quantify the settle time, which was found to be approximately five hours. The temperature dependence experiment was repeated on a settled sample and the results from this match that seen in the full study. The fitted parameters were determined with a larger error, this was due to the lower degrees in the freedom effecting the fitting process. It has been shown there is a trade-off between accuracy and speed when using the DOE methodologies. The use of the DOE models would not have been possible if the added accuracy and repeatability of the automation software had not been in place.

12.1 The impact of e-Science on the second harmonic generation experiment

In developing the environmental monitoring and experimental control systems the data collected from the SHG experiment has been greatly improved. The introduction of the additional software controlled devices, such as the auto-diluter, has allowed new experiments to be completed to a high degree of accuracy. The use of the message broker has allowed these new components to be integrated to the experimental system with minimal changes to the software.

In the past, work on the SHG rig required a researcher to manually rotate the input half-wave plate up to 90 times for each output polarisation selected; in the case of 90 rotations the accuracy must be within a degree. Similarly, when running adsorption isotherms the researcher would manually dilute the sample. This significantly increases the time taken to complete an experiment and therefore it may span over multiple days. The 30^{o}C adsorption isotherm of PNP has shown how changing samples between runs can have a detrimental effect on the collected results. Through introducing the automation of the equipment a huge amount of error can be reduced, or even removed. This is achieved by both providing highly accurate motors to position the components of the system and through removal of human error.

This also improves the repeatability of the experiments. Each experiment is defined through an XML message sent to the message broker and is stored automatically into the experimental database. As the storage of the results and experimental data is part of the automated system, the experiment is easily repeatable with exact experimental parameters. The parameters are available for review through the SHG results website. The use of the electronically controlled components also mean when repeating experiments the accuracy will be increased.

The reduction in time required to complete a SHG experiment has been the largest improvement. A full polarisation scan experiment can be completed over approximately

three hours, reduced from several days in some cases. As the system is automated the researcher can either analyse previous data or begin preparing for the next experiment while the first is running; in the past this would have not been possible. This has allowed all experiments reported in this work to be completed in the final year of the research. Previous doctoral research projects have focused on the study of a single parameter from only one compound.

The experimental process has also been optimised using DOE methodologies; this has only been made possible by the automation and control software. This has lead to the discovery of novel chemical properties, the formation of the PNP surface, which would not have been possible to observe using previous experimental techniques.

The environmental data acquisition software has increased the trust in the data as well as providing a tool to determine potential causes in failed experiments. The modular design approach, intercommunicating via the message broker, has allowed for rapid development and integration of 'plug-in' components to the core acquisition system. As the results to the given experiment are automatically available through the SHG website, the researcher can instantly identify potential environmental causes for anomalous results.

The use of the on-line blog as an ELN has increased the volume of data recorded as well as improving its quality. As it is on-line, it is available to access from anywhere across the globe; this has made the sharing of results and ideas between distributed researchers simple. The integration of a data repository with the blog framework has allowed results supplementary to the SHG results to be link together, making the data storage and processing significantly easier.

12.2 Further work

Many aspects of the project could be investigated further to fully understand and be exploited to full potential as discussed in the individual chapters. A full test of the system could be carried out by providing the tools developed in this project to another research group. Observation of how easily the deployment and usage is will give an indication of the success of the software. The software was developed in a modular 'plug-and-play' format to allow easy integration of new components, the success of this can be monitored in seeing what components other users develop and integrate into their own deployment.

List of Publications

1. Greenhalgh, C., Glover, K., Humble, J., Robinson, J., Wilson, S., Frey, J., Page, K. and De Roure, D. (2008) Combining System Introspection with User-Provided Description to Support Configuration and Understanding of Pervasive systems. In: Third International Conference on Pervasive Computing and Application (ICPCA2008), October 2008, Alexandria, Egypt

2. Wilson, S., Frey, J., *The smartLab: Experimental and environmental control and monitoring of the chemistry laboratory*, cts, pp.85-90, 2009 International Symposium on Collaborative Technologies and Systems, 2009

3. Wilson, S., Milsted, A., Frey, J., *Comment by sketch: A Picture Says a Million Words*, Microsoft eScience Workshop, pp.20-22, Microsoft eSceince Workshop 2009, 2009

4. Wilson, S., Coles, S., Frey, J., *Second Life: The Next Virtual Laboratory?*, All Hands Meeting Oxford, pp.48-49, All Hands Meeting 2009, 2009

5. Wilson, S., Frey, J., *Critical Zone Observatories and Sensor Repositories*, All Hands Meeting Oxford, pp.11, All Hands Meeting 2009, 2009

6. Coles, S. and Wilson, S. (2010) Second Life: the next virtual laboratory? In, International Conference on Education and New Learning Technologies - EduLearn10, Barcelona, ES 05 - 07 Jul 2010. , 7pp.

7. Wilson, S., Frey, J., *The Impact of eScience on Real Science: the second harmonic generation experiment*, All Hands Meeting Cardiff, All Hands Meeting 2010, 2010

Bibliography

[1] M Atkinson. e-Science. [Web page] http://www.rcuk.ac.uk/escience/. [Accessed 28th September 2010].

[2] Jeremy G. Frey. Comb-e-chem - an e-science research project. In Martyn Ford, David Livingstone, John Dearden, and Han Van der Waterbeemd, editors, *EuroQSAR 2002 Designing Drugs and Crop Protectants: processes, problems and solutions*, pages 395–398. Blackwell, Oxford, UK, 2003.

[3] What is e-Science? E-science Grid. [web page] http://www.esciencegrid.org.uk/what-e-science.html. [Accessed 27th September 2010].

[4] oreChem. Microsoft Research. [web page] http://research.microsoft.com/en-us/projects/orechem/. [Accessed 25th October 2010].

[5] Jamie M. Robinson, Jeremy G. Frey, Andy J. Stanford-Clark, Andrew D. Reynolds, and Bharat V. Bedi. Chemistry by mobile phone (or how to justify more time at the bar). In *UK e-Science All Hands Meeting 2005*, UK, September 2005.

[6] SQL*LIMS. LabVantage Solutions. [web page] http://www.sqllims.com/. [Accessed 28th September 2010].

[7] BioRails. The Edge. [web page] http://www.biorails.com/. [Accessed 28th September 2010].

[8] Office 2010. Microsoft TechNet Library. [web page] http://technet.microsoft.com/en-gb/library. [Accessed 21st Decemeber 2010].

[9] LimsLink. Labtronics. [web page] http://www.labtronics.com/lims_interfacing.htm. [Accessed 28th September 2010].

[10] m.c. schraefel, Gareth Hughes, Hugo Mills, Graham Smith, Terry Payne, and Jeremy Frey. Breaking the book: Translating the chemistry lab book into a pervasive computing lab environment. 2004.

[11] R Scoffin. The new wave in electronic laboratory notebook systems. *Chemical Biology & Drug Design*, 67(2):184–185, FEB 2006.

[12] Mark C. Fishman and Jeffery A. Porter. Pharmaceuticals: A new grammar for drug discovery. *Nature*, 437(7058):491–493, September 2005.

[13] Public Record Office. Retention scheduling - health and safety records, 1998.

[14] David J Drake. Eln implementation challenges. *Drug Discov Today*, 12:647–649, 2007.

[15] Nexxis. Labtronics. [web page] http://www.labtronics.com/electronic_laboratory_notebook.htm. [Accessed 28th September 2010].

[16] E-WorkBook Suite. IDBS. [web page] http://www.idbs.com/ELN/. [Accessed 28th September 2010].

[17] SAFE. SAFE-BioPharma Association. [web page] http://www.safe-biopharma.org/. [Accessed 28th September 2010].

[18] CambridgeSoft Life Science Enterprice Solutions. CambridgeSoft. [web page] http://www.cambridgesoft.com/. [Accessed 28th September 2010].

[19] LabTrove. University of Southampton. [web page] http://www.labtrove.org/. [Accessed 28th September 2010].

[20] Arthur W. Adamson and Alice P. Gast. *Physical Chemistry of Surfaces.* Wiley Interscience, 1997.

[21] Kenneth B. Eisenthal. Liquid interfaces. *Accounts of Chemical Research*, 26(12):636–643, 1993.

[22] K B Eisenthal. Equilibrium and dynamic processes at interfaces by second harmonic and sum frequency generation. *Annual Review of Physical Chemistry*, 43(1):627–661, 1992.

[23] Duncan J Shaw. *Introduction to Colloid and Surface Chemistry.* Butterworth-Heinemann, 2003.

[24] Robert W. Boyd. *Nonlinear Optics.* Academic Press, 1992.

[25] K. B. Eisenthal. Liquid interfaces probed by second-harmonic and sum-frequency spectroscopy. *Chemical Reviews*, 96:1343–1360, 1996.

[26] Hiroyuki Fujiwara. *Spectroscopic ellipsometry: Principles and applications.* Wiley & Sons, Ltd, 2007.

[27] Jillian C Wallis, C L Borgman, Matthew S Mayernik, Alberto Pepe, Nithya Ramanathan, and Mark Hansen. Know thy sensor: Trust, data quality, and data integrity in scientific digital libraries. 2007.

[28] E. Snchez, A. B. Portas, J. Vega, J. M. Agudo, K. J. McCarthy, M. Ruiz, E. Barrera, and S. Lpez. Autonomous acquisition systems for tj-ii: controlling instrumentation with a fourth generation language. *Fusion Engineering and Design*, 71(1-4):123 – 127, 2004. 4th IAEA Technical Meeting on Control, Data Acquistion, and Remote Participation for Fusion Research.

[29] LabVIEW. National Instruments. [web page] http://www.ni.com/labview/. [Accessed 18th August 2010].

[30] VNC remote control software. RealVNC. [web page] http://www.realvnc.com/. [Accessed 18th August 2010].

[31] W.B. Heinzelman, A.L. Murphy, H.S. Carvalho, and M.A. Perillo. Middleware to support sensor network applications. *Network, IEEE*, 18(1):6 – 14, jan/feb 2004.

[32] Kay Römer, Oliver Kasten, and Friedemann Mattern. Middleware challenges for wireless sensor networks. *SIGMOBILE Mob. Comput. Commun. Rev.*, 6:59–61, October 2002.

[33] FJA Cardoso. A universal system for laboratory data acquisition and control. *IEEE Transactions On Nuclear Science*, 47(2, Part 1):154–157, APR 2000.

[34] D. Comer. *Internetworking with TCP/IP: Principles, protocols, and architecture.* Prentice Hall, 2006.

[35] J.F. Wang, Peter W. Tse, L.S. He, and Ricky W. Yeung. Remote sensing, diagnosis and collaborative maintenance with web-enabled virtual instruments and miniservers. *The International Journal of Advanced Manufacturing Technology*, 24:764–772, 2004. 10.1007/s00170-003-1753-4.

[36] A. Ramakrishnan. 16-bit embedded web server. *Sensors for Industry Conference, 2004. Proceedings the ISA/IEEE*, pages 187 – 193, 2004.

[37] Tao Lin, Hai Zhao, Jiyong Wang, Guangjie Han, and Jindong Wang. An embedded web server for equipment. pages 345 – 350, may. 2004.

[38] I. Klimchynski. Extensible embedded web server architecture for internet-based data acquisition and control. *Sensors Journal, IEEE*, 6(3):804 –811, jun. 2006.

[39] IBM United Kingdom. IBM. [web page] http://www.ibm.com/uk/. [Accessed 12th January 2009].

[40] MQ - Telemetry Transport. MQTT.org. [web page] http://mqtt.org/. [Accessed 12th January 2009].

[41] Microsoft MSDN. Publish/Subscribe. [web page] http://msdn.microsoft.com/en-us/library/ff649664.aspx. [Accessed 6th September 2010].

[42] Integrating monitoring and telemetry devices as part of enterprise information resources. IBM. [web page] http://www-306.ibm.com/software/integration/mqfamily/integrator/telemetry/pdfs/telemetry_integration_ws.pdf. [Accessed 13th June 2008].

[43] Message Broker Help. IBM. [web page] http://publib.boulder.ibm.com/infocenter/wmbhelp/v6r0m0/index.jsp?topic=/com.ibm.etools.mft.doc/ac10840_.htm. [Accessed 13th June 2008].

[44] What's Middleware. ObjectWeb. [web page] http://middleware.objectweb.org/. [Accessed 5th January 2011].

[45] S. Wilson. Middleware infrastructure for remote experimental monitoring and control. Master's thesis, University of Southampton, 2007.

[46] The Perl directory - Perl org. http://www.perl.org/. [web page] http://www.perl.org/. [Accessed 17th July 2008].

[47] Labjack – USB/Ethernet based data acquisition and control. Labjack. [web page] http://www.labjack.com/. [Accessed 12th January 2008].

[48] GNU Screen. GNU Prjoect. [web page] http://www.gnu.org/software/screen/. [Accessed 17th August 2010].

[49] A J G Hey and A E Trefethen. The data deluge: An e-science perspective, 2003.

[50] C L. Borgman, Jillian C. Wallis, Matthew S. Mayernik, and Alberto Pepe. Drowning in data: Digital library architecture to support scientific use of embedded sensor networks. In *International Conference on Digital Libraries*, pages 269–277, 2007.

[51] MySQL :: The world's most popular open source database. MySQL. [web page] http://www.mysql.com/. [Accessed 23rd July 2008].

[52] Welcome to the OGC website. OGC. [Web page] http://www.opengeospatial.org/. [Accessed 12th July 2008].

[53] Sensor Web Enablement WG. OGC. [Web page] http://www.opengeospatial.org/projects/groups/sensorweb. [Accessed 21st Decemeber 2010].

[54] Sensor Observation Service. OGC. [Web page] http://www.opengeospatial.org/standards/sos. [Accessed 21th December 2010].

[55] A. Na and M. Priest. Sensor observation service - ogc 06-009r6, October 2007.

[56] Welcome to 52North. 52North.org. [Web page] http://52north.org/. [Accessed 12th January 2009].

[57] Apache Tomcat. The Apache Software Foundation. [web page] http://tomcat.apache.org/. [Accessed 5th January 2011].

[58] PostgreSQL: The world's most advanced open source database. PostgreSQL. [web page] http://www.postgresql.org/. [Accessed 16th June 2008].

[59] International Organisation of Standardisation. Information technology – Open Systems Interconnection – Remote Procedure Call (RPC). ISO/IEC 11578:1996.

[60] RRDtool: logging & graphing. RRDtool. [web page] http://oss.oetiker.ch/rrdtool/. [Accessed 24th July 2008].

[61] Pachube :: connecting environments, patching the planet. Pachube. [web page] http://www.pachube.com/. [Accessed 7th September 2010].

[62] Extended Environments Markup Language: EEML. EEML. [web page] http://www.eeml.org/. [Accessed 7th September 2010].

[63] Apache. The Apache HTTP Server Project. [web page] http://httpd.apache.org/. [Accessed 24th July 2008].

[64] Simile — Timeplot. Simile. [web page] http://simile.mit.edu/timeplot/. [Accessed 12th January 2009].

[65] SensorBase — Sharing is caring. SensorBase. [web page] http://sensorbase.org/. [Accessed 9th September 2010].

[66] SensorPedia. SensorPedia. [web page] http://www.sensorpedia.org/. [Accessed 9th September 2010].

[67] M.J. Cox and J.E.F. Baruch. Robotic telescopes: An interactive exhibit on the world-wide web. In *The second International Conference of the World-Wide-Web*, 1994.

[68] R.M. Parkin, C.A. Czarnecki, R. Safaric, and D.W. Calkin. A pid servo control system experiment conducted remotely via internet. *Mechatronics*, 11:833–843, July 2002.

[69] A. Neto, H. Fernandes, A. Duarte, B. B. Carvalho, J. Sousa, D. F. Varcarcel, M. Hron, and C. A. F. Varandas. Firesignal - data acquisition and control. *Fusion Engineering And Design*, 82(5-14):1359–1364, OCT 2007.

[70] Ian M. Atkinson, Douglas du Boulay, Clinton Chee, Kenneth Chiu, Paul Coddington, Andrea Gerson, Tristan King, Donald F. McMullen, Romain Quilici, Peter Turner, Andrew Wendelborn, Mathew Wyatt, and Donglai Zhang. Developing cima-based cyberinfrastructure for remote access to scientific instruments and collaborative e-research. In *Proceedings of the fifth Australasian symposium on ACSW frontiers - Volume 68*, ACSW '07, pages 3–10, Darlinghurst, Australia, Australia, 2007. Australian Computer Society, Inc.

[71] V. J. Harward, J. A. del Alamo, V. S. Choudhary, and K. deLong. ilab: A scalable architecture for sharing online experiments. In *Proceedings of the ICEE 2004 Conference*, 2004.

[72] National Instruments. G Programming Reference Manual, January 1998.

[73] Randall Bramley, Kenneth Chiu, Tharaka Devadithya, Nisha Gupta, Charles Hart, John C. Huffman, Kianosh Huffman, Yu Ma, and Donald F. McMullen. Instrument monitoring, data sharing, and archiving using common instrument middleware architecture (cima). *Journal of Chemical Information and Modeling*, 46(3):1017–1025, 2006.

[74] GridCC. GridCC. [web page] http://www.gridcc.org. [Accessed 10th September 2010].

[75] F. Davoli, G. Spano, S. Vignola, and S. Zappatore. Labnet: Towards remote laboratories with unified access. *Instrumentation and Measurement, IEEE Transactions on*, 55(5):1551 –1558, oct. 2006.

[76] M. Cooper, A. Donnelly, and J. Ferreira. Remote controlled experiments for teaching over the internet: A comparison of approaches developed in the pearl project. In *Proceedings of the ASCILITE Conference 2002*, 2002.

[77] I. Foster, H. Kishimoto, and A. Savva. The Open Grid Services Architecture, Version 1.5, July 2006.

[78] The Common Component Architecture Forum. CCA. [Accessed 13th September 2010].

[79] EIA/TIA. Interface between data terminal equipment and data communication equipment employing serial data interchange, August 1969.

[80] iUI. Google Code. [web page] http://code.google.com/p/iui/. [Accessed 5th October 2010].

[81] Linden Research, Inc. Second Life. [web page] http://secondlife.com/. [Accessed 5th October 2010].

[82] Erik Christensen, Francisco Curbera, Greg Meredith, and Sanjiva Weerawarana. Web Services Description Language (WSDL) 1.1, March 2001.

[83] Simile — Timeplot. Simile. [web page] http://www.simile-widgets.org/timeplot. [Accessed 5th January 2011].

[84] PHP: Hypertext Preprocessor. PHP. [web page] http://www.php.net/. [Accessed 24th July 2008].

[85] Simile — Timeline. Simile. [web page] http://www.simile-widgets.org/timeline/. [Accessed 13th September 2010].

[86] JSON. JSON. [web page] http://www.json.org/. [Accessed 21st December 2010].

[87] Simile — Exhibit. Simile. [web page] http://simile-widgets.org/exhibit/. [Accessed 13th September 2010].

[88] Google Chart Tools / Image Charts. Google Code. [web page] http://code.google.com/apis/chart/. [Accessed 13th September 2010].

[89] The R Project for Statistical Computing. R project. [web page] http://www.r-project.org/. [Accessed 27th July 2010].

[90] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich. Generation of optical harmonics. *Phys. Rev. Lett.*, 7(4):118–119, Aug 1961.

[91] Victor Mizrahi and J.E. Sipe. Phenomenological treatment of surface second-harmonic generation. *J. Opt. Soc. Am.*, 5:660–667, 1988.

[92] B. Dick, A. Gierulski, G. Marowsky, and G. A. Reider. Determination of the nonlinear optical susceptibility $\chi^{(2)}$ of surface layers by sum and difference frequency generation in reflection and transmission. *Applied Physics B: Lasers and Optics*, 38:107–116, October 1985.

[93] Levi M. Haupert and Garth J. Simpson. Chirality in nonlinear optics. *Annual Review of Physical Chemistry*, 60(1):345–365, 2009. PMID: 19046125.

[94] Pierre-Francois Brevet. *Surface Second Harmonic Generation.* PU Polytechniques, 1996.

[95] Wen-kai Zhang, Hong-fei Wang, and De-sheng Zheng. Quantitative measurement and interpretation of optical second harmonic generation from molecular interfaces. *Phys. Chem. Chem. Phys.*, 8:4041–4052, 2006.

[96] Esther R Rousay. *Studies at the Solution/Air Interface by Experiment and Molecular Dynamics Simulation.* PhD thesis, School of Chemistry, University of Southampton, 2005.

[97] Richard I. Masel. *Principles of Adsorption and Reaction on Solid Surfaces.* Wiley Interscience, 1996.

[98] Lefteris Danos. Non-linear spectroscopic studies at interfaces: Experiment and theory. Master's thesis, Department of Chemistry, University of Southampton, 2003.

[99] H Sahsah, S Djendli, and J Monin. A new method of birefringence measurements using a faraday modulator. application to measurements of stress-optical coefficients. *Measurement Science and Technology*, 11(3):N46, 2000.

[100] T D Wolfenden, P E G Baird, J A Deeny, and M Irie. Use of a faraday modulator in a laser polarimeter for optical forward-wave level-crossing experiments in atomic vapours. *Measurement Science and Technology*, 1(10):1060, 1990.

[101] E.W. Washburn. International critical tables of numerical data, physics, chemistry and technology (1st electronic edition).

[102] S. Haslam. *Second Harmonic Generation at the liquid / air interface*. PhD thesis, Faculty of Science, 1999.

[103] A. Fordyce. *Second Harmonic Generation at the liquid / air interface*. PhD thesis, Faculty of Science, 1999.

[104] Andrew John Timson. *Second Harmonic Generation of Molecules Located at the Air / Water Interface*. PhD thesis, School of Chemistry, University of Southampton, 2000.

[105] Alfio A. Tamburello Luca, Philippe Hebert, Pierre F. Brevet, and Hubert H. Girault. Surface second-harmonic generation at air/solvent and solvent/solvent interfaces. *J. Chem. Soc., Faraday Trans.*, 91:1763–1768, 1995.

[106] M. Cynthia Goh and Kenneth B. Eisenthal. The energetics of orientation at the liquid-vapor interface of water. *Chemical Physics Letters*, 157(1-2):101 – 104, 1989.

[107] V P Sokhan and D J Tildesley. The free surface of water: molecular orientation, surface potential and nonlinear susceptibility. *Molecular Physics*, 92(4):625–640, 1997.

[108] Rodolphe Antoine, Francois Bianchi, Pierre F. Brevet, and Hubert H. Girault. Studies of water/alcohol and air/alcohol interfaces by second harmonic generation. *J. Chem. Soc., Faraday Trans.*, 93:3833–3838, 1997.

[109] R. Greef and J. G. Frey. The water-like film on water. *Phys. Status Solidi (c)*, 5(5):1184–1186, 2008.

[110] Andrew J. Bell, Jeremy G. Frey, and Ted J. VanderNoot. Second harmonic generation by p-nitrophenol at water/air and water/heptane interfaces. *J. Chem. Soc., Faraday Trans.*, 88:2027–2030, 1992.

[111] Christian Leuenberger, Mary P. Ligocki, and James F. Pankow. Trace organic compounds in rain. 4. identities, concentrations, and scavenging mechanisms for phenols in urban air and rain. *Environmental Science & Technology*, 19(11):1053–1058, 1985.

[112] Jens Lttke and Karsten Levsen. Phase partitioning of phenol and nitrophenols in clouds. *Atmospheric Environment*, 31(16):2649 – 2655, 1997. The Great Dun Fell Cloud Experiment 1993, Eurotrac sub-project Ground-based Cloud Experiment (GCE).

[113] Mark A.J. Harrison, Silvia Barra, Daniele Borghesi, Davide Vione, Cecilia Arsene, and Romeo Iulian Olariu. Nitrated phenols in the atmosphere: a review. *Atmospheric Environment*, 39(2):231 – 248, 2005.

[114] Maria Paluch and Maria Filek. Effect of para-substituted phenols on the surface potential and on the surface tension at the water/air interface. *Journal of Colloid and Interface Science*, 73(1):282 – 286, 1980.

[115] Alfio A. Tamburello-Luca, Philippe Hebert, Pierre F. Brevet, and Hubert H. Girault. Resonant-surface second-harmonic generation studies of phenol derivatives at air/water and hexane/water interfaces. *J. Chem. Soc., Faraday Trans.*, 92:3079–3085, 1996.

[116] Daniel A. Higgins, Michael B. Abrams, Shannon K. Byerly, and Robert M. Corn. Resonant second harmonic generation studies of p-nitrophenol adsorption at condensed-phase interfaces. *Langmuir*, 8(8):1994–2000, 1992.

[117] Kankan Bhattacharyya, E. V. Sitzmann, and K. B. Eisenthal. Study of chemical reactions by surface second harmonic generation: p-nitrophenol at the air–water interface. *The Journal of Chemical Physics*, 87(2):1442–1443, 1987.

[118] S. Haslam, S. G. Croucher, C. G. Hickman, and J. G. Frey. Surface second harmonic generation studies of the dodecane/water interface: the equilibrium and kinetic behaviour of p-nitrophenol and tri-n-butyl phosphate. *Phys. Chem. Chem. Phys.*, 2:3235–3245, 2000.

[119] Alonso Castro, Shaowei Ong, and Kenneth B. Eisenthal. Studies of molecular properties at the surface of a liquid jet by second harmonic generation. *Chemical Physics Letters*, 163(4-5):412 – 416, 1989.

[120] Kaustuv Das, Nilmoni Sarkar, Swati Das, Anindya Datta, Debnarayan Nath, and Kankan Bhattacharyya. Salt effect on p-nitrophenol at the water surface: a surface second harmonic generation study. *J. Chem. Soc., Faraday Trans.*, 92:4993–4996, 1996.

[121] William H. Steel and Robert A. Walker. Solvent polarity at an aqueous/alkane interface: the effect of solute identity. *Journal of the American Chemical Society*, 125(5):1132–1133, 2003. PMID: 12553791.

[122] Nilmoni Sarkar, Kaustuv Das, Swati Das, Debnarayan Nath, and Kankan Bhattacharyya. Effect of urea and surfactant on p-nitrophenol at the water surface: a surface second-harmonic generation study. *J. Chem. Soc., Faraday Trans.*, 91:1769–1773, 1995.

[123] George W. Gokel, W. Matthew Leevy, and Michelle E. Weber. Crown ethers: Sensors for ions and molecular scaffolds for materials and biological models. *Chemical Reviews*, 104(5):2723–2750, 2004. PMID: 15137805.

[124] Wladyslaw Walkowiak, Sang Ihn Kang, Lewis Stewart, Grace Ndip, and Richard A. Bartsch. Effect of ring size variation within lipophilic crown ether carboxylic acids on the selectivity and efficiency of competitive alkali-metal cation solvent extraction into chloroform. *Analytical Chemistry*, 62(18):2022–2026, 1990.

[125] Saeid Azizian, Kaoru Kashimoto, Takashi Matsuda, Hiroki Matsubara, Takanori Takiue, and Makoto Aratono. Interfacial tension studies of crown ethers at air/water and hexane/water interfaces. *Journal of Colloid and Interface Science*, 316(1):25 – 30, 2007.

[126] Esther R. Rousay, Hongchen Fu, Jamie M. Robinson, Jonathan W. Essex, and Jeremy G. Frey. Grid-based dynamic electronic publication: A case study using combined experiment and simulation studies of crown ethers at the air/water interface. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 363(1833):pp. 2075–2095, 2005.

[127] Adinela Cazacu, Christine Tong, Arie van der Lee, Thomas M. Fyles, and Mihail Barboiu. Columnar self-assembled ureido crown ethers: an example of ion-channel organization in lipid bilayers. *Journal of the American Chemical Society*, 128(29):9541–9548, 2006. PMID: 16848492.

[128] Marcin Dulak, Rmi Bergougnant, Katharina M. Fromm, Hans R. Hagemann, Adeline Y. Robin, and Tomasz A. Wesolowski. Water trapped in dibenzo-18-crown-6: Theoretical and spectroscopic (ir, raman) studies. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 64(2):532 – 548, 2006.

[129] Katharina M. Fromm, Rmi D. Bergougnant, and Adeline Y. Robin. Di-benzo-18-crown-6 and its derivatives as ligands in the search for ion channels. *Zeitschrift fr anorganische und allgemeine Chemie*, 632(5):828–836, 2006.

[130] Petru Niga, Wendy King, Jonas Hedberg, C. Magnus Johnson, Jeremy G. Frey, and Mark W. Rutland. Crown ethers at the aqueous solution-air interface: 1. assignments and surface spectroscopy. *Phys. Chem. Chem. Phys.*, 13:7930–7938, 2011.

[131] Petru Niga, C. Magnus Johnson, Jeremy G. Frey, and Mark W. Rutland. Crown ethers at the aqueous solution-air interface. part 2. electrolyte effects, ethylene oxide hydration and temperature behaviour. *Phys. Chem. Chem. Phys.*, 13:7939–7947, 2011.

[132] A. Male. Comparison of the behaviour of c18 model compounds and natural sea slicks using langmuir-blodgett films and ellipsometry. Master's thesis, Chemistry, 2009.

[133] P. Niga. *Self Assembly at the Liquid Air Interface*. PhD thesis, School of Chemical Science and Engineering, 2010.

[134] Peter Atkins and Julio de Paula. *Atkins' Physical Chemistry*. Oxford, 7th edition, 2002.

[135] H. D. Flack and G. Bernardinelli. The mirror of galadriel: looking at chiral and achiral crystal structures. *Crystal Engineering*, 6(4):213 – 223, 2003. Patterns in Nature.

[136] A. Papakostas, A. Potts, D. M. Bagnall, S. L. Prosvirnin, H. J. Coles, and N. I. Zheludev. Optical manifestations of planar chirality. *Phys. Rev. Lett.*, 90(10):107404, Mar 2003.

[137] J. Miones, P. Dynarowicz-Latka, J. Miones, J. M. Rodriguez Patino, and E. Iribarnegaray. Orientational changes in dipalmitoyl phosphatidyl glycerol langmuir monolayers. *Journal of Colloid and Interface Science*, 265(2):380 – 385, 2003.

[138] Amlia M. Gonalves da Silva, Rute S. Romo, Ana Lucero Caro, and Juan M. RodrI[combining acute accent]guez Patino. Memory effects on the interfacial characteristics of dioctadecyldimethylammonium bromide monolayers at the air-water interface. *Journal of Colloid and Interface Science*, 270(2):417 – 425, 2004.

[139] G. Box, W. G. Hunter, and J. S. Hunter. *Statistics for Experimenters*. Wiley, 1978.

[140] C. J. Wu and M. Hamanda. *Experients: Planning, Analysis, and Parameter Design Optimization*. Wiley Interscience, 2000.

[141] A. Forrester. Design and analysis of experiments in engineering design. 2008.

[142] Martin R. Owen, Chris Luscombe, Lai, Sonya Godbert, Derek L. Crookes, and David Emiabata-Smith. Efficiency by design: Optimisation in process research. *Organic Process Research & Development*, 5(3):308–323, 2001.

[143] Stefanie Biedermann and David C. Woods. Optimal designs for generalised nonlinear models with application to second harmonic generation experiments. In *S3RI Methodology Working Papers*, 2009.

Appendix

Support data DVD

A data DVD has been made available containing a number of appendices to support this thesis.

Appendix A – Environmental monitoring and experimental control scripts

The scripts written to handle the data acquisition and experimental control via the message broker have been made available for further development, re-purposing and review. Each script is contained within a directory relating to the acquisition process it handles or the equipment it controls.

Appendix B – Experimental results, fitting scripts and fitting results

As discussed in the appropriate chapters, a large number of datasets have been generated from a combination of experiments. These datasets have undergone a number of analytical processes to produce the final results. This analysis was carried out using R[89] scripts. Each script not just produces an output file containing the determined parameters but a number of subsidiary plots to verify and support the determined parameters. These datasets, scripts and output from those scripts have been made available for further analysis and review. The directory structure is divided into compounds, within each compound directory are directories relating to the parameter under investigation. Within this directory is the dataset, any supporting data and the processing scripts. The output from these scripts are located within the output sub directory of each folder. The table below details the mapping from tables in the thesis to files within the appendix.

Table number	Location
7.2	Water/Output/anova.csv
7.3	Water/Output/RDwithErrors.csv
8.6	PNP/TemperatureDependence/Output/temp_values.csv
8.7	PNP/TemperatureDependence/Output/temp_values.csv
8.8	PNP/TemperatureDependence/Output/T_fit_temp_values.csv
8.10	PNP/ConcentrationDependence/Output/conc_values.csv
8.12	PNP/AdsorptionIsotherm/Output/langmuir_values.csv
8.13	PNP/TemperatureDependence/Output/temp_ratios.csv
8.14	PNP/TemperatureDependence/Output/RDwithErrors.csv
8.15	PNP/TemperatureDependence/Output/RDwithErrors.csv
8.16	PNP/DRFTemperatureDependence/Output/RDF_values.csv
8.18	PNP/ConcentrationDependence/Output/conc_ratios.csv
8.19	PNP/ConcentrationDependence/Output/RDwithErrors.csv
8.20	PNP/ConcentrationDependence/Output/RDwithErrors.csv
8.21	PNP/AdsorptionIsotherm/Output/langmuir_values.csv
8.23	PNP/SurfaceTension/Output/frumkin_values.csv
8.24	PNP/SurfaceTension/Output/langmuir_values.csv
8.26	PNP/RemoveWater/Output/RDvalues.csv
9.2	NB15C5/TemperatureDependence/Output/temp_values.csv & temp_values_omit.csv
9.4	NB15C5/SaltDependence/Output/salt_values_omit.csv
9.6	NB15C5/AdsorptionIsotherm/Output/langmuir_values.csv
9.7	NB15C5/TemperatureDependence/Output/temp_ratios_omit.csv
9.8	NB15C5/TemperatureDependence/Output/RDwithErrors_omit.csv & RDwithErrors_average.csv
9.9	NB15C5/TemperatureDependence/Output/RDwithErrors_omit.csv & RDwithErrors_average.csv
9.10	NB15C5/TemperatureDependence/Output/RDwithErrors_omit.csv & RDwithErrors.csv
9.11	NB15C5/TemperatureDependence/Output/RDwithErrors_omit.csv & RDwithErrors.csv
9.12	NB15C5/DRFTemperatureDependence/Output/compare_theory.csv
9.13	NB15C5/SaltDependence/Output/temp_ratios_omit.csv

9.14	NB15C5/ConcentrationDependence/Output/conc_ratios.csv
9.15	NB15C5/ConcentrationDependence/Output/RDwithErrors.csv
9.16	NB15C5/ConcentrationDependence/Output/RDwithErrors.csv
9.17	NB15C5/SurfaceTension/Output/NB15C5-surfacetension.csv
10.2	B15C5/Anova/Output/anova_values.csv
10.3	B15C5/TemperatureDependence/Output/temp_values.csv
10.5	B15C5/SaltDependence/Output/salt_values.csv
10.7	B15C5/TemperatureDependence/Output/temp_ratios.csv
10.8	B15C5/SaltDependence/Output/salt_ratios.csv
10.9	B15C5/LithiumAnova/Output/anova_values.csv
10.10	B15C5/SurfaceTensionTime/B15C5-surfacetension_settle.csv
10.11	B15C5/SurfaceTensionSettled/B15C5-surfacetension.csv
10.12	B15C5/SurfaceTensionSettled/B15C5-surfacetension.csv
11.4	DOE/TempDepAfterSettle/temp_ratios.csv

Appendix C – Second Harmonic generation experimental results database

An export of the SHG experimental database has been made available for further data analysis. The file can be imported into an existing MySQL installation to recreate that of the one created during this work.

Appendix D – Electronic laboratory notebook copy relating to this work

An ELN was used to record the work, thoughts, meetings, experiments and other ramblings completed during the PhD. The laboratory notebook used was one developed in-house, LabTrove[19]. This blog based laboratory notebook solution allows the user to download a static download of the notebook in its present form. This download has been presented in this appendix.