THE HONG KONG
POLYTECHNIC UNIVERSITY
香港理工大學
Pao Yue-kong Library
包玉剛圖書館

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

Pao Yue-kong Library, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

http://www.lib.polyu.edu.hk

This thesis in electronic version is provided to the Library by the author. In the case where its contents is different from the printed version, the printed version shall prevail.

DEVELOPMENT OF NEW INHIBITORS FOR THE BACTERIAL GLYCOSYLTRANSFERASE: COMPUTATIONAL DOCKING, SYNTHESIS AND BIOASSAYS

WANG YONG

Ph.D

The Hong Kong Polytechnic University

2014

The Hong Kong Polytechnic University

Department of Applied Biology and Chemical Technology

Development of New Inhibitors for the Bacterial Glycosyltransferase: Computational Docking, Synthesis and Bioassays

WANG Yong

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

January, 2014

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____

WANG Yong

January, 2014

Abstract

Bacterial drug resistance caused by abuse and misuse of antibiotics represents a widespread problem in our modern society. There is an urgent need to discover new and novel antibacterial agents with new targets, otherwise the drug resistant bacteria will become a great threat to human health. The bacterial glycoltransferase (GT), which plays a vital role in the biosynthesis of cell wall using lipid II as substrate, has potential to become a novel antibacterial target as the sequence of its encoding gene is highly conserved in both wild-type and drug-resistant strains. Currently, there is no clinical drug targeting at this enzyme. The only known glycosyltransferase inhibitor, moenomycin, is used as a growth promoter in animal feed due to its poor bioavailability. After all, discovery of novel glycosyltransferase inhibitors is a challenging but rewarding research area.

Computer-aided drug design (CADD) has been widely used in the discovery of new drugs against specific binding targets. This technique can simulate the binding process between receptors and ligands, and the binding affinity could be estimated by score function. Compared to the traditional method of drug development, CADD accelerates discovery speed, saves time, reduces cost and improves hit-rate.

To discover new antibacterial agents with glycosyltransferase as the target, over 3,000,000 compounds were computationally screened against the GT domain of *S. aureus* penicillin binding protein 2 (PDB: 2OLV, 2.8Å) by Internal Coordinate Mechanics (ICM) software. A hit compound with final score -34.5, 2-(3-(2-carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(m-tolyl) acetamide (compound **GT10**) was found to possess a weak antibacterial ability to *S. aureus* and *B. subtilis* (MIC, 192 μg mL^{-1}). Saturation-transfer difference (STD)-NMR confirmed the interaction between **GT10** and the GT domain of *S. aureus*.

In order to optimize the structure of **GT10**, analogues of the hit compound with the basic isatin core were designed and successfully synthesized with acceptable yields. The structures of all these derivatives were confirmed by ^{1}H-NMR and mass spectrometry.

The structure activity relationship was rationalized according to the MIC results of these **GT10** analogues against *S. aureus*, *B. subtilis* and *E. coli*. The results revealed that substituents such as nitro, fluorine or ether at the *meta*-position on the phenyl ring of the N-phenyl amide group can improve the antibacterial ability. Besides, the guanidyl group attached to the isatin core is essential to the activity. To better understand the binding between GT domain and **GT10** analogues, a docking simulation between **GT10-22, GT10-27** and *S. aureus* GT domain was performed by ICM. This model

shows that the phenyl group in the N-phenyl amide group and the isatin core establish hydrophobic interaction with Ile195, Val233 and Pro234 whereas the guanidyl group enhances binding by forming three H-bonds with Lys155 and Ser160 of the GT domain. Interaction between the phenyl rings of **GT10-27** and the GT domain was confirmed by STD-NMR.

Acknowledgements

I truly express my gratitude to my supervisor Prof. K. Y. Wong for his support in all aspects during my whole PhD study period. I am greatly inspired by his valuable suggestions and comments regarding my research project. It is really grateful that he teaches me to think holistically and provides me such a challenging but significant project.

Prof. Y. C. Leung is deeply appreciated for giving me valuable advice on the design of biological assays. Also, I need to thank Dr. F. Y. Chan for teaching me computational docking techniques, Dr. W. H. Chung for teaching me molecular cloning, Dr. S. C. Yan for the assistance in STD-NMR experiments, and Dr P. K. So for the mass spectra measurements of proteins.

I am also deeply grateful to all my colleagues, specially, Dr. W. L. Wong, Dr. Ning Sun, and Mr. H. K. Lui. They are greatly thanked for their encouragement and support in the last three years.

Especially, from the bottom of my heart, I wish to express my utmost gratitude to my parents, my wife and my family. It is truly great of what

they have done for supporting and encouraging me to go for higher education.

Last but not least, I would like to acknowledge the student stipends administrated by the Research Committee of the Hong Kong Polytechnic University from 2010 to 2013.

Table of Contents

Certificate of Originality **I**

Abstract **II**

Acknowledgements **V**

Table of Contents **VII**

List of Abbreviations **XI**

Chapter 1 Introduction **1**

1.1 Bacterial infections and antibacterial drugs 2

1.2 Antibacterial drug resistance 4

1.2.1 Causes of drug resistance 4

1.2.2 Mechanism of antibiotic drug resistance 4

1.3 Peptidoglycan biosynthesis 8

1.3.1 Cell wall structure 8

1.3.2 Biosynthetic pathway of peptidoglycan 10

1.4 GT as a promising antibacterial drugs target 12

1.5 Bacterial transglycosylation inhibitors 14

1.5.1 Natural transglycosylation inhibitors 14

1.5.1.1 Glycopeptide antibiotics 14

1.5.1.2 Lantibiotics 18

1.5.1.3 Moenomycin family 19

1.5.2 Potential GT inhibitors 22

1.6 Computer-aided drug design 24

1.6.1 Introduction 24

1.6.2 General principles in computer-aided drug design 26

1.6.2.1 Molecular mechanics 26

1.6.2.2 Molecular dynamics 27

1.6.2.3 Score functions 27

1.6.3 Virtual screening in drug discovery 28

1.6.3.1 Structure-based virtual screening 28

1.6.3.2 Ligand-based virtual screening 31

1.7 Saturation-transfer difference nuclear magnetic resonance 32

1.8 Aims and objectives 35

Chapter 2 Identification of hit compounds by virtual high throughput screening and bioassays 36

2.1 Introduction 37

2.2 Experimental 39

2.2.1 Materials 39

2.2.2 Virtual screening based on molecular docking 42

2.2.3 Antibacterial test 47

2.2.4 Preparation of recombinant *S.aureus* GT domain 48

2.2.5 Saturation-transfer difference (STD)-NMR 53

2.3 Results and discussion 54

2.3.1 Structure-based virtual screening of potential GT inhibitors 54

2.3.1.1 Redocking of the known inhibitor moenomycin A 54

2.3.1.2 *In silico* screening and compounds selection 56

2.3.2 Antibacterial activity of hit compounds obtained from virtual screening 61

2.3.3 Expression of *S. aureus* GT domain 66

2.3.4 Saturation-transfer difference (STD)-NMR spectra 70

2.4 Concluding remarks 72

Chapter 3 Synthesis and characterization of GT10 and its analogues 73

3.1 Introduction 74

3.2 Design of **GT10** analogues 78

3.3 Experimental 79

3.4 Results and discussion 82

3.4.1 Product yields and characterization 82

3.4.2 Synthetic routes for **GT10** and its analogues 91

3.5 Concluding remarks 93

Chapter 4 Structure activity relationship study and molecular docking simulation 94

4.1 Introduction 95

4.2 Experimental 96

4.3 Results and discussion 98

4.3.1 Structure activity relationship study 98

4.3.2 Molecular docking simulation 105

4.3.3 Saturation-transfer difference (STD)-NMR 107

4.3.4 Cytotoxicity 111

4.4 Concluding remarks 113

Chapter 5 Conclusions **114**

Appendix I: ^{1}H-NMR spectra of GT10 and its analogues **118**

Appendix II: MS spectra of GT10 and its analogues **149**

References **180**

List of Abbreviations

AMBER	Assisted Model Building with Energy Refinement
ATCC	American Type Culture Collection
BPMC	Biased probability Monte Carlo
B. subtilis	*Bacillus subtilis*
BSA	Bovine serum albumin
CADD	Computer-aided drug design
CH_3CH_2OH	Ethanol
CH_2Cl_2	Dichloromethane
CH_3CN	Acetonitrile
CH_3COOH	Acetic acid
Da	Dalton
D-Ala-D-Ala	D-alanine-D-alanine
D-Ala-D-Lac	D-alanine-D-lactate
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
GA	Genetic algorithm
GT	Glycosyltransferase
HTS	High-throughput screening
ICM	Internal coordinate mechanics

IPTG	Isopropyl β-D-1-thiogalactopyranoside
K_2CO_3	Potassium carbonate
LB medium	Luria-Bertani medium
MIC	Minimum inhibitory concentration
MMFF	Merck Molecular Force Field
MoeA	Moenomycin A
MRSA	Methicillin-resistant *Staphylococcus aureus*
nm	nanometer
OD_x	Optical density at wavelength of x nm
PBP	Penicillin binding protein
PDB	Protein data bank
RMSD	Root mean square deviation
RNA	Ribonucleic acid
rpm	Revolution per minute
S. aureus	*Staphylococcus aureus*
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
STD-NMR	Saturation-transfer difference nuclear magnetic resonance
TLC	Thin layer chromatography
UV	Ultraviolet
VRE	Vancomycin resistant *enterococci*

Chapter 1

Introduction

1.1 Bacterial infection and antibacterial drugs

As a matter of fact, most of the bacteria are harmless. Some bacteria, such as *bifidobacteria* in the gastrointestinal tract of human are even beneficial to human health as they can improve the intake of calcium, reduce cholesterol, stimulate the immune response and slow down the aging process [1-3]. However, pathogenic bacteria such as *Mycobacterium tuberculosis*, *Streptococcus* and *Pseudomonas* can cause very serious health problems in humans. Every year, millions of people die from tuberculosis caused by *Mycobacterium tuberculosis* [4]. Pneumonia is still a leading infectious disease in children and the aged [5-6].

Antibacterial drugs, which kill or suppress the growth of bacteria, can be classified into several groups according to the mechanism of action. The mechanisms include interference of cell wall biosynthesis (penicillin), DNA and RNA synthesis (quinolones, rifampin), cell membrane synthesis (polymyxin), essential protein synthesis (aminoglycosides, chloramphenicol) and folate synthesis (sulfonamides) (Figure 1.1).

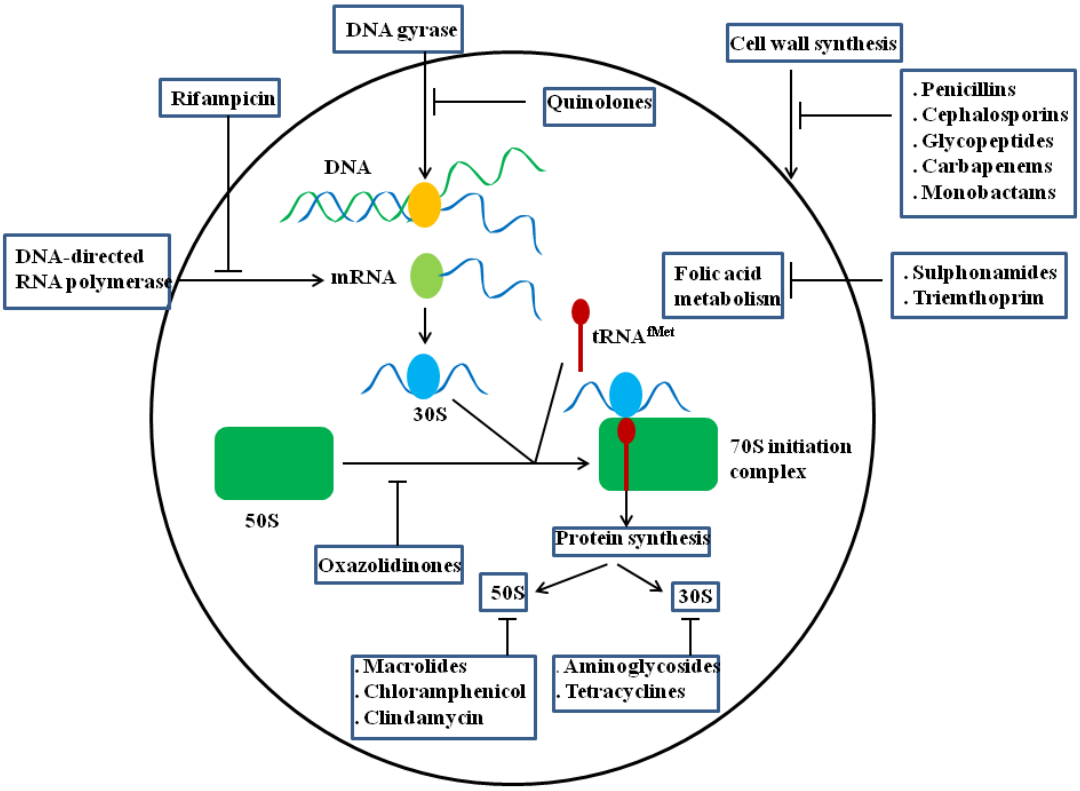


Figure 1.1 Targets of antibacterial agents. Adopted from reference [7].

1.2 Antibiotic drug resistance

1.2.1 Causes of drug resistance

Antibiotics have been used for many years, and these drugs definitely contribute significantly to save human life. However, bacteria which survive the treatment of antibacterial drugs are drug-resistant species. Even more seriously, some of these species such as Methicillin-resistant *Staphylococcus aureus* (MRSA) are resistant to more than one type of drugs [8-9]. Multiple drug resistance, which is a growing phenomenon in our society, will dramatically threaten the health of human being. To some extent, the appearance of drug resistance is an outcome of evolution by natural selection. It is a strategy to adapt to the inconstant and formidable natural environment [10]. Nevertheless, the misuse and abuse of antibacterial drugs by human badly accelerates this evolution [11-12]. Besides, the mutation of drug resistance can be transferred in bacteria [13].

1.2.2 Mechanism of antibiotic drug resistance

The mechanism of resistance to antibiotics drug can be summarized into four major pathways as shown in Figure 1.2.

I. Modification of drug targets.

This is the most common mechanism. The structure of binding proteins will be modified by bacteria, so that antibacterial drugs cannot recognize or bind to them, leading to a diminution of drug effect. Take vancomycin as an example: in vancomycin-resistant *Enterococci* (VRE), the bacterial cells can

produce D-alanine-D-lactate terminus instead of D-alanine-D-alanine on the cell wall precursor, which vancomycin cannot bind to. This change decreases the binding affinity of vancomycin to the receptor dramatically [14].

II. Decrease of drug uptake

In this pathway, bacteria will alter the penetration of drug to its membrane, so that antibiotics cannot reach the target or do not accumulate enough to the effective concentration. In addition, bacterial cells can also express proteins that act as pumps on the surface of cell membrane. With efflux of antibiotics by these pumps, the concentration of drugs inside the bacterial cell is kept at a relatively low level. Quinolone resistance is mediated by an endogenous system that transports the drug out of bacteria [15].

III. Deactivation of antibiotics by enzymatic modification

This mechanism aims to destroy the chemical structure of antibiotics. The β-lactam antibiotics, such as penicillin G and cefalexin, kill bacteria by inhibiting the formation of cell wall. Bacterial cells can develop resistance by producing β-lactamases which hydrolyze the β-lactam ring and hence inactivate the antibiotic [16]. The chemical structure of aminoglycosides can also be modified by bacterial enzymes, so that they lose the ability of interrupting protein synthesis due to the lowering of binding affinity to RNA [17].

IV. Metabolic bypass of the inhibition reaction

This strategy is uncommon in bacterial resistance to antibiotics. Under this mechanism bacteria can survive by undertaking the inhibition of drugs using alternative enzyme which also binds to the antibiotic [18].

As antibiotic resistant pathogens will exhibit severe threat to humans, the development of novel antibacterial agents aiming at new targets is of great urgency.

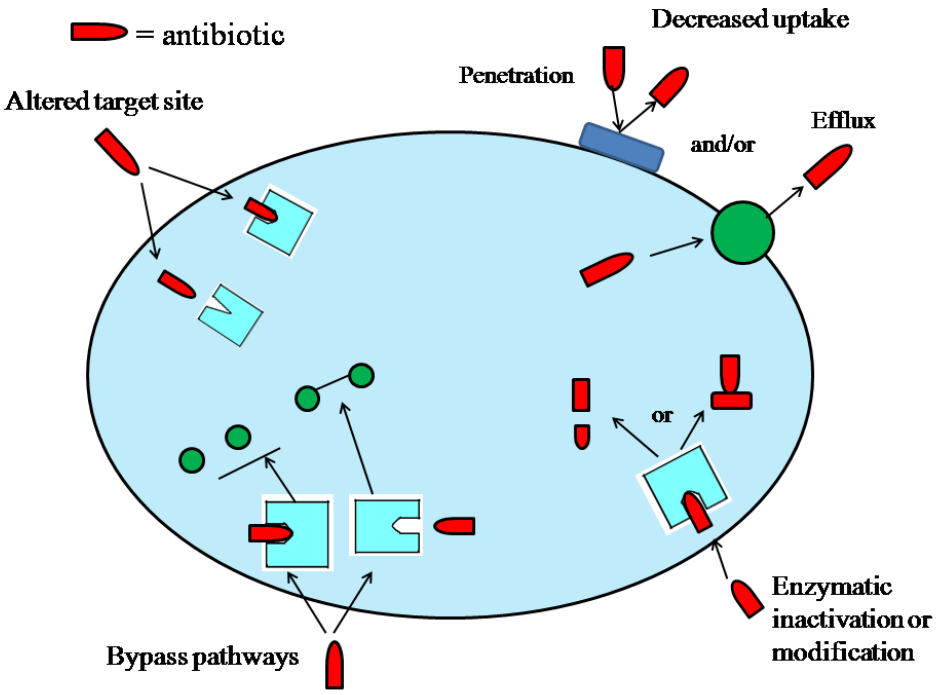


Figure 1.2 The four major pathways of antibiotic resistance mechanism. Adopted from reference [18].

1.3 Peptidoglycan biosynthesis

1.3.1 Cell wall structure

Cell wall, which is tough and flexible, plays an extraordinarily vital role in the living of bacteria. The major function of cell wall is to act like a shield protecting bacteria from the high internal osmotic pressure and the nasty external living environment. Peptidoglycan is the basic component of bacterial cell wall (Figure 1.3). Different bacterial species may have different cell wall structure. Compared to Gram-positive bacteria, the cell wall of Gram-negative bacteria contains an extra outer membrane. However, the composition of peptidoglycan polymer is similar in all bacteria and its biosynthesis is highly conserved [19].

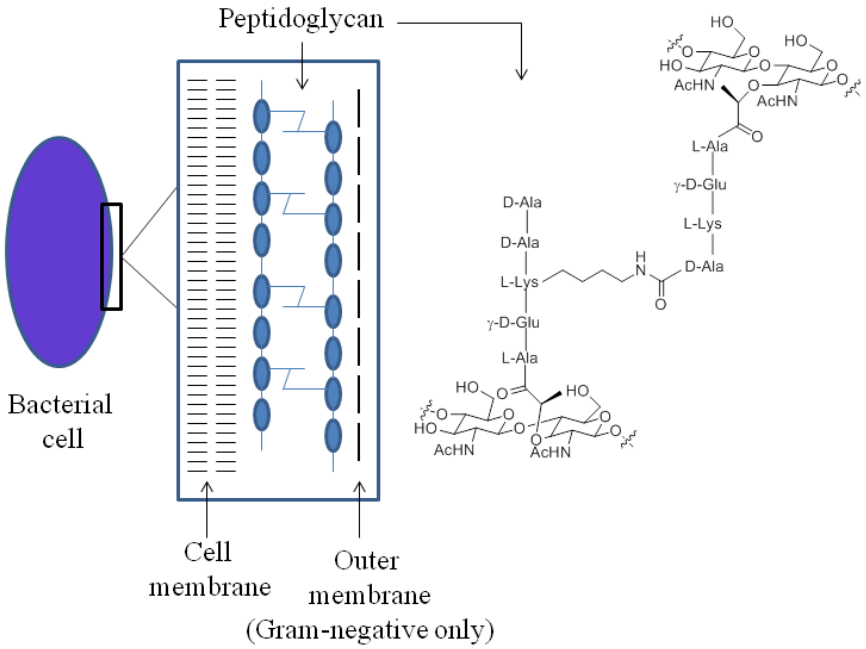


Figure 1.3 Structure of bacterial cell wall. Adopted from reference [19].

1.3.2 Biosynthetic pathway of peptidoglycan

Peptidoglycan, which consists of sugars and amino acids, forms a mesh-like layer outside the plasma membrane of bacteria. Bacterial peptidoglycan synthesis is initiated in the cytoplasm with the synthesis of UDP-N-acetyl-muramyl-pentapeptide (UDP-MurNAc-pentapeptide), which is then transferred via MraY to a C55 undecaprenol phosphate carrier to produce lipid I. Subsequently, MurG transfers N-acetylglucosamine (NAG) from UDP-N-acetylglucosamine (UDP-GlcNAc) to lipid I to generate lipid II, which is then polymerized into peptidoglycan by glycosyltransferase (GT) and transpeptidase (TP). The sugar component consists of alternating residues of β-(1, 4) linked N-acetylglucosamine (NAG) and N-acetylmuramic (NAM) acid residues. A peptide chain consisted of 3 to 5 amino acids is attached to the N-acetylmuramic acid. The peptide chains are finally cross-linked to form the 3D mesh-like layer [20-22].

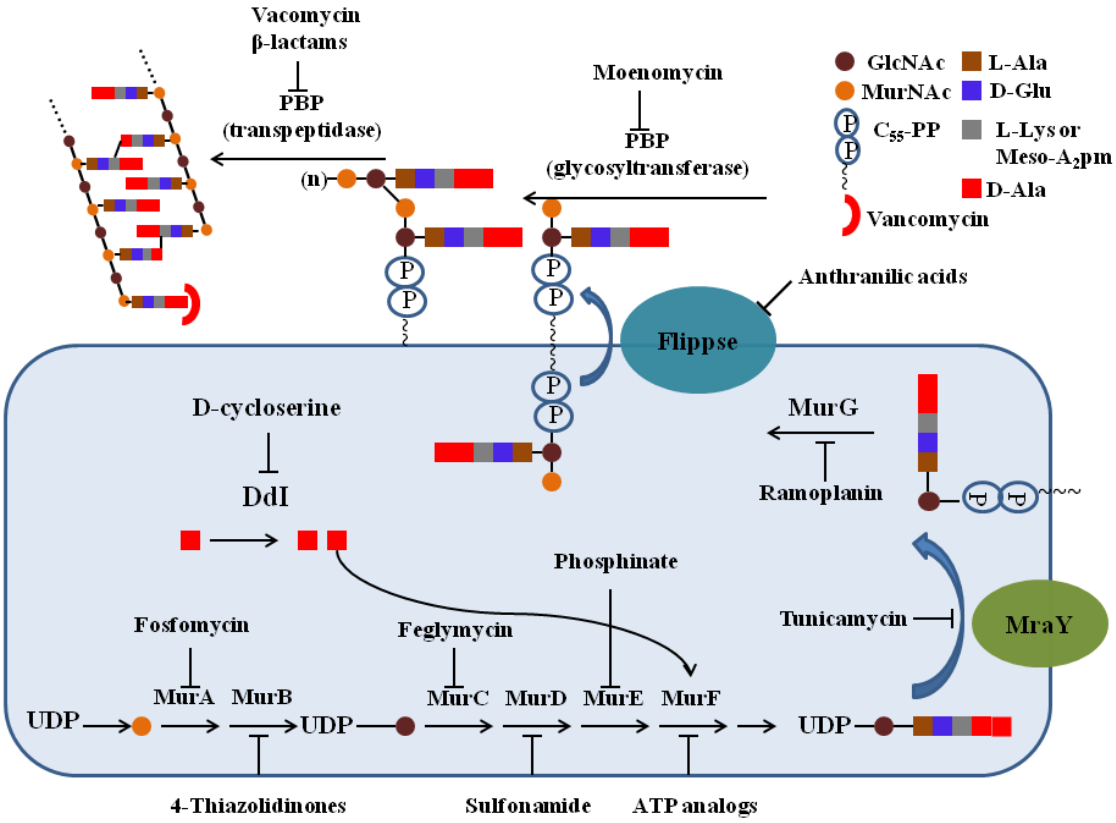


Figure 1.4 Biosynthesis of peptidoglycan and the targets of some inhibitors.

Adopted from reference [23].

1.4 GT as a promising antibacterial drugs target

As discussed before, glycosyltransferase (GT) is involved in the biosynthesis of peptidoglycan. This enzyme catalyzes the incorporation of lipid II into glycan strands, which are then crosslinked by transpeptidase (TP) to form peptidoglycan. Usually, monofunctional GTs are not essential to bacterial cell wall synthesis and the indispensable ones exist as one domain in the penicillin binding proteins (PBPs), such as PBP 1b and PBP 2 [24]. These PBPs can be classified into two groups, namely, high-molecular-weight and low-molecular-weight PBPs [25]. The bifunctional high-molecular-weight PBPs are essential for bacterial survival as they catalyze the formation of cell wall and maintain the cellular structure in bacteria [26-27]. Transpeptidase, the other domain of high-molecular-weight PBPs, is the target of the classical antibacterial drug β-lactams. Yet, a disturbingly high degree of drug resistance against β-lactams has been developed over the past decades [28].

In recent years, GT has aroused great enthusiasm from scientists to consider it as a novel target for discovery of antibacterial drugs. The enzyme possesses several advantages to be a good drug target candidate. First of all, GT is an essential enzyme in the biosynthesis of cell wall, which provides a stable living environment for bacteria. In addition, it is easy for inhibitors to access, as they are located on the outside of bacterial membrane [29]. Moreover, the structure of GT is highly conserved no matter in wild-type or drug resistant strains. Multiple sequence alignment indicated that there are

five similar motifs of GT in different types of bacteria, even in some drug resistant species [30-31]. Finally, GT has no counterpart in the eukaryotic cells [32].

With much effort from scientists, the gene encoding GT can now be cloned into specific plasmid and this protein has been successfully expressed in host cells. It is more exciting that the crystal structure of some vital high-molecular weight PBPs such as PBP 1b and PBP 2 are known now [33-34]. The detailed structural information on the GT domain provided by these crystal structures guides the design of new GT inhibitors. Furthermore, there is no antibiotic directly targeted GT in the clinical trial phase until now. All the information above indicates that GT is an attractive and promising target for discovery of novel antibacterial drugs.

1.5 Bacterial transglycosylation inhibitors

1.5.1 Natural transglycosylation inhibitors

In the transglycosylation step, lipid IIs are polymerized into glycan chains by glycosyltransferase. Natural products that inhibit the transglycosylation step can be divided into two groups, namely, those that bind to the substrate lipid II and those that inhibit the enzyme glycosyltransferase.

1.5.1.1 Glycopeptides antibiotics

The basic structures of this category of antibacterial drugs are composed of glycosylated cyclic or polycyclic nonribosomal peptides (Figure 1.5). The mechanistic action of these compounds is to form hydrogen bonds through the peptide with the D-alanine-D-alanine end of lipid II. On one hand, this action prevents the transfer of lipid II to the surface of the bacterial outer membrane as the glycopeptides cannot penetrate the cytoplasm [35]. On the other hand, the binding of peptide to lipid II also inhibits the transpeptidation step [36-37]. However, after several decades of use of glycopeptide antibiotics, some bacteria have developed resistance to these drugs by modifying the D-Ala-D-Ala to D-Ala-D-Lac, which decreases the binding affinity by almost 1000 folds (Figure 1.6) [38-39].

Besides, researchers have proved that some N-alkylated vancomycin, such as chlorobiphenyl-vancomycin, exhibit activity on the VRE [40]. The mechanism of action of these novel compounds is attributed to the direct inhibition of transglycosylation instead of binding to D-Ala-D-Ala

[41].These findings provide a new direction for discovery of novel antibacterial agents. Vancomycin is considered as the antibiotic of last resort. It will be disastrous if its resistance is transferred to other drug resistant pathogens, such as MRSA.

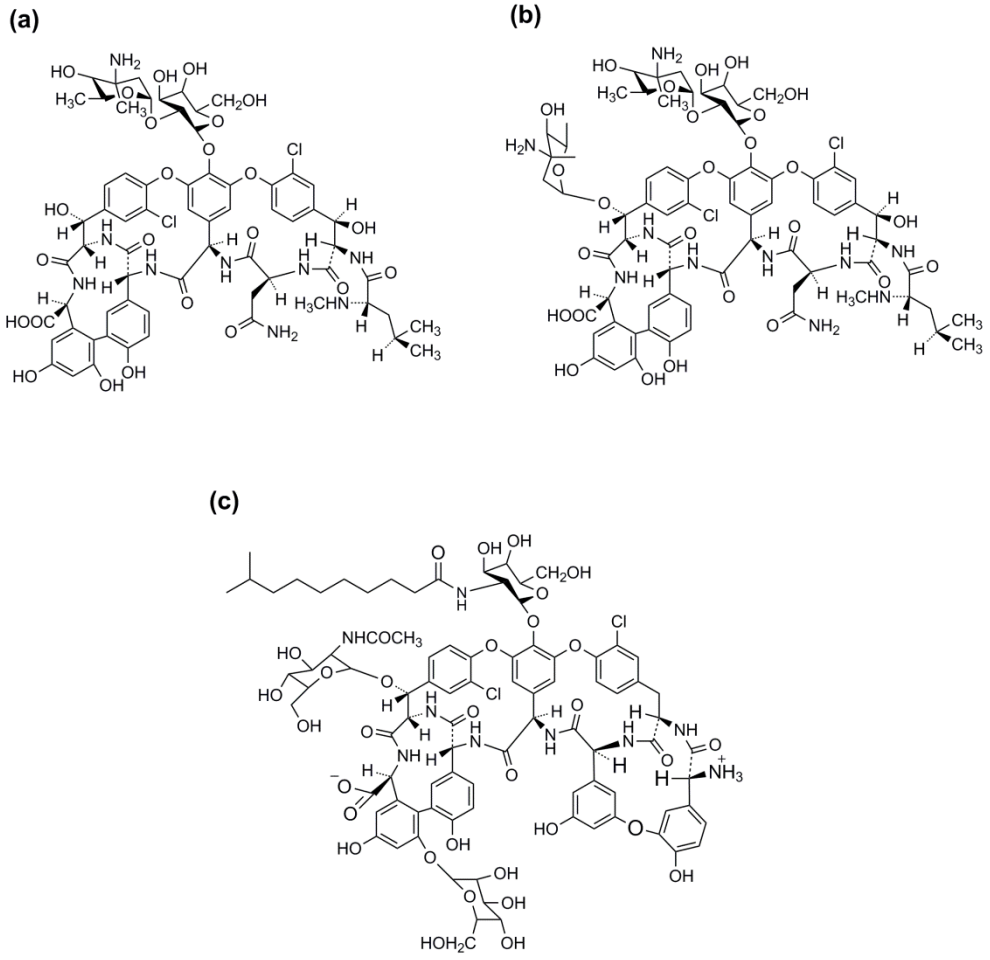


Figure 1.5 Glycopeptides antibiotics. (a) Vancomycin, (b) Chloroeremomycin, (c) Teicoplanin. [42]

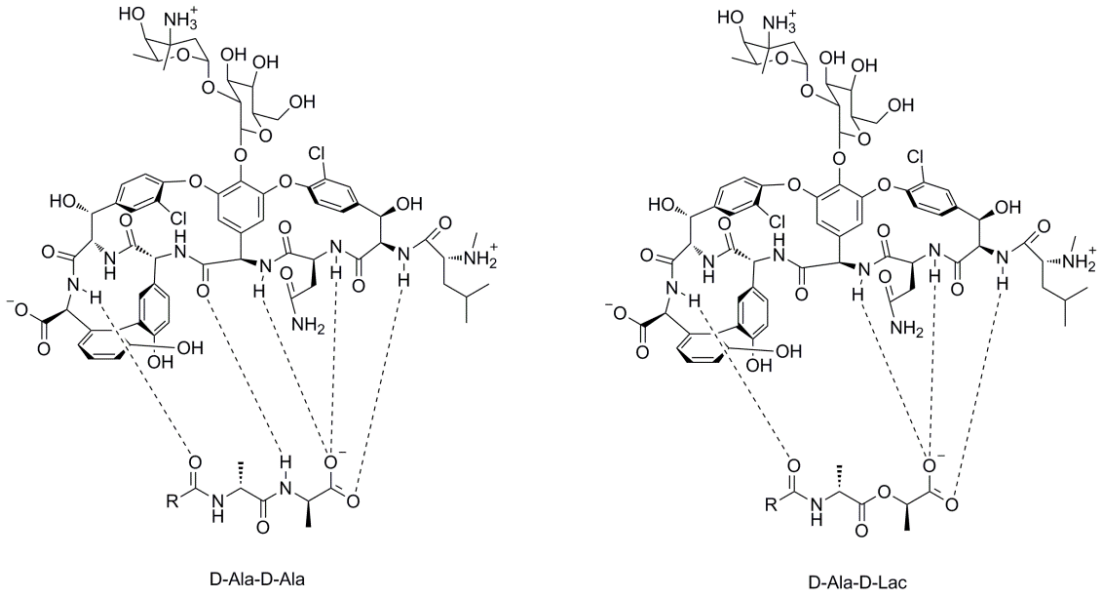


Figure 1.6 Mechanism of action and resistance of vancomycin [43]. The binding between vancomycin and D-Ala-D-Ala forms five H-bonds in wild type strains. Replacement of D-Ala by D-Lac in vancomycin-resistant strains causes decrease of one H-bond.

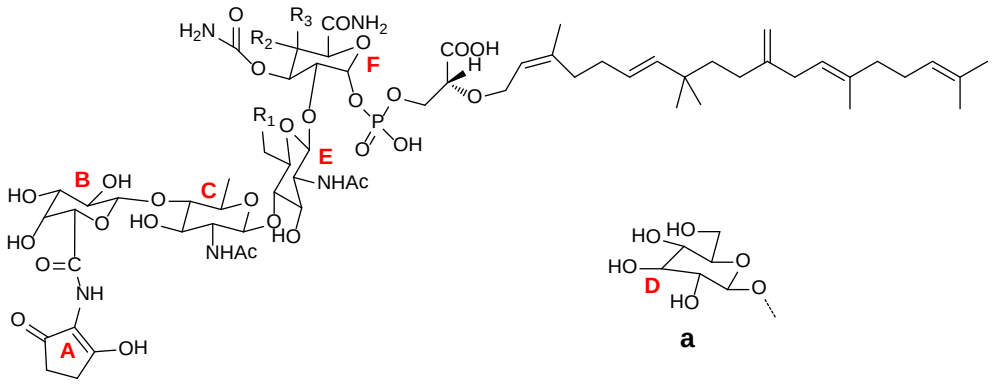
1.5.1.2 Lantibiotics

Lantibiotics, metabolites of certain gram-positive bacteria, are peptide antibiotics composed of unusual amino acids, such as lanthionine, and contain several disulfide bridges [44]. In general, this kind of antibiotics can be classified into two groups according to the structure and mechanism of action. Type A lantibiotics, e.g. nisin, are cationic and amphipathic molecules which kill bacteria by forming pores on the bacterial membrane [45]. They have broad antibacterial spectrum and low cytotoxicity. In contrast, the structure of type B lantibiotics, mersacidin and actagardine, is globular and they can interfere with the transglycosylation by forming a unique complex with lipid II [46]. This mechanism is quite different with that of vancomycin, which binds to the D-Ala-D-Ala tail of lipid II, since their activity to VRE and vancomycin susceptible *Enterococci* are almost identical [47-48].

1.5.1.3 Moenomycin family

Moenomycins, natural products from *Streptomyces*, are the only known class of natural compounds that inhibit the transglycosylation in the formation of bacterial cell wall [49]. Among this family, Moenomycin A (MoeA) is most abundant and its structure consists of a pentasaccharide with a C_{25} isoprene (Figure 1.7).

Although MoeA exhibits antibacterial ability to many species, especially Gram-positive bacteria, it is merely used as a growth promoter in animal feed under the trademark Flavomycin for more than 30 years and no high-level resistance has been detected until now [50]. MoeA has a long half-life and shows some toxicity upon parenteral administration, but it is not effective in human due to poor absorption properties. Nevertheless, it is believed that further research on the interaction between MoeA and GTs can lead to the development of new antibacterial agents based on these compounds.



	R_1	R_2	R_3
Moenomycin A	**a**	CH3	OH
Moenomycin A_{12}	**a**	OH	H
Moenomycin C_4	OH	CH3	OH
Moenomycin C_3	H	CH3	OH
Moenomycin C_1	H	OH	H

Figure 1.7 Main members of the moenomycin family.

The structure of MoeA is complicated, despite the fact that its total synthesis has been successfully completed [51]. Over the last two decades, numerous works, including degradation studies and analogues synthesis, have been performed to evaluate the structure-activity relationship of MoeA [52-57]. The results manifest that the A ring is not essential for interaction, since the activity only slightly decreases after removing it. The trisaccharide C-E-F and the disaccharide E-F, the minimal pharmacophore structures, can both inhibit GT reaction *in vitro*, but only the trisaccharide structure possesses good MIC results. Besides, the C_{25} lipid moiety is also indispensible. Removing this part or reducing the length of aliphatic chain invalidates the antibacterial activity. Moreover, it has been testified that this chain will anchor into the cytoplasmic membrane, which is in favor of selective binding between MoeA and GT [37, 58].

The binding pose between MoeA and PBP 2 was obtained by cocrystallization [33]. The MoeA-bound structure shows that the MoeA molecule mimics the lipid II in the GT pocket. The sugar rings B to F are in a chair conformation that form a twisted plane and ring F blocks the putative Glu114 catalytic residue. The phosphoric acid diester group is located under this plane and is in such an orientation that it potentially directs the C_{25} lipid group toward the cytoplasmic membrane. This model is consistent with the results of structure-activity relationship [56].

1.5.2 Potential GT inhibitors

In order to discover novel glycosyltransferase inhibitors, researchers have concentrated on the modification of MoeA and the GT substrate lipid II with interesting results. However, until now there is not much progress on small molecules that can inhibit GT enzymatic activity. Several compounds obtained by fluorescence polarization high throughput screening are considered to possess this potential [59-60] (Figure 1.8). Yet, none of them is in clinical use. Finding small molecules as GT inhibitors is still in a very challenging but promising research direction.

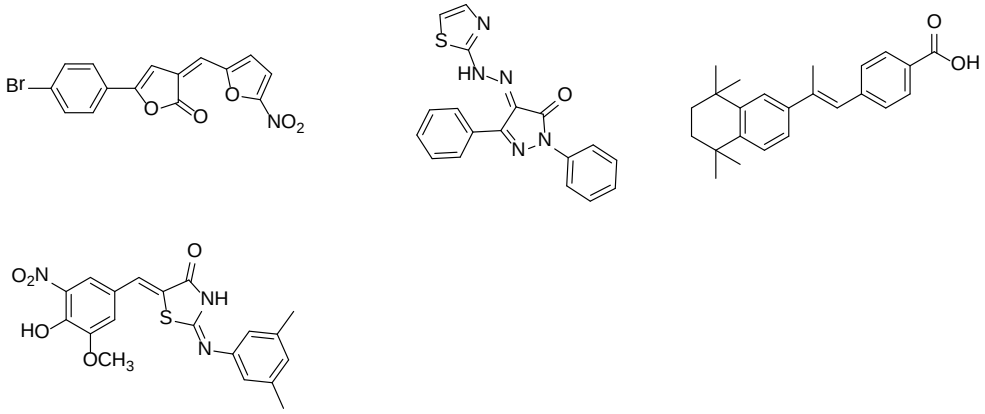


Figure 1.8 Some potential GT inhibitors obtained by fluorescence polarization high throughput screening [59-60].

1.6 Computer-aided drug design

1.6.1 Introduction

Computer-aid drug design (CADD) is the use of computational technique to simulate the interaction between small molecules and specific biological macromolecules, usually enzymes. It predicts the binding affinity and pose through evaluation of docking results [61]. The hit compounds will then be testified by a series of experiments. All these obtained information will guide the design and optimization of lead compounds. In the first instance, CADD was used to design compounds that bind to human haemoglobin [62]. Although CADD only has a history of about 30 years, it is growing rapidly. With the details of more protein crystal structures known, and databases containing millions of chemical structures built, CADD now plays a vital role in the search of lead compounds against desired drug targets [63].

CADD can overcome the time-consuming process of searching drug leads by trial and error in the traditional drug development, which has very low hit-rate [64-66]. Thus, CADD can reduce cost and save time with better hit-rate [67-68]. The general procedure for CADD is summarized in Figure 1.9.

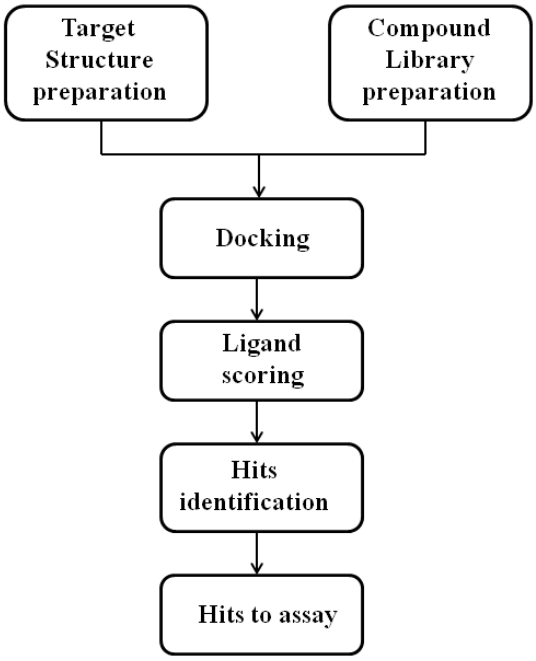


Figure 1.9 Typical procedures for drug screening by CADD. Adopted from reference [69].

1.6.2 General principles in computer-aided drug design

1.6.2.1 Molecular mechanics

Molecular mechanics is a method based on classical Newton mechanics to calculate the molecular energy in different conformations by force field. The potential energy of a molecular system is the sum of covalent bonded energy and noncovalent bonded energy. The former is the energy from bond stretching, angle bending and torsion rotation. The latter is generated by electrostatic and van der Waals interactions. Unbound energy is more significant in computational chemistry, since it represents most interactions between molecules [70]. At present, there are many types of force fields to calculate the energy of different molecular system. Take Assisted Model Building with Energy Refinement (AMBER) for example, it is a widely used force field to handle macromolecules, such as proteins [71]. Merck Molecular Force Field 94 (MMFF94) is commonly applied in the treatment of small molecules [72].

One of the important applications of molecular mechanics is to find the lowest energy conformation of a molecular system by a proper algorithm such as steepest descents, conjugate gradients and Newton-Raphson method. This would be used in molecular docking and lead optimization. The energy values calculated from molecular mechanics have no actual meaning as absolute quantities, and they are only useful in comparing the different conformations of the same structure.

1.6.2.2 Molecular dynamics

The interaction between molecules is not a static process. Molecular dynamics, based on Newton's law of motion, is widely used to simulate the physical movements of molecular systems and provides the details of interaction at atomic level. This technique is always applied in the refinement of protein crystal structures [73]. Besides, in molecular docking, it can make ligand and protein structures flexible in the docking process. By doing this, the interaction in docking is closer to the real situation and the software can generate more accurate predicted complexes and binding energy [74]. Monte Carlo simulation is another method to generate random conformations of small molecules [75]. Both methods play significant role in molecular docking.

1.6.2.3 Score functions

Score function is a tool to predict the binding affinity and orientation between two molecules (usually receptors and ligands) after the completion of docking. They are classified into three categories, namely force field based, empirical based and knowledge-based [76]. Force field based score functions use the van der Waals and electrostatic interaction between molecules in the docking process. Also, internal energy such as bond stretching, angle and torsions is included [77]. Empirical based score functions consider that the interaction energy could be divided into different components. The protein-ligand complexes are calibrated with multiple

linear regressions using experiment data [78-79]. The knowledge-based score function predicts the binding between molecules using the information obtained from known high resolution X-ray crystal structures of protein-ligand complexes as references [80-81]. These three score functions are widely used in all kinds of docking programs.

1.6.3 Virtual screening in drug discovery

Virtual screening is a technique using docking programs to evaluate libraries which contain millions of small molecule structures against certain specific targets (usually proteins). The binding affinity between small molecules and target receptor is then evaluated through score functions [82]. This process accelerates the discovery of drugs, as usually a smaller number of structures need to be synthesized and further tested. Virtual screening is generally classified into structure-based and ligand-based methods.

1.6.3.1 Structure-based virtual screening

Structure-based virtual screening, which has been widely used in the early stage of drug discovery, involves the docking of tens of thousands of ligands into a specific pocket of protein. The docking program will predict ligand orientation and binding affinity. High score compounds are considered as leads. In this method, high resolution three dimensional coordinates of proteins usually obtained from X-ray diffraction and the availability of chemical libraries are essential [83].

Usually, structure-based virtual screening adopts the flexible ligands against rigid receptor. However, most proteins are also flexible in solution and have various conformations especially after introduction of ligands [84]. So more accurate results can be obtained by using flexible receptor, though this process is much more complicated and time-consuming [85].

Several methods to construct flexible receptor are now available, which employ the ensemble of multiple crystal structures, generation of flexible side chains and simulation by molecular dynamics [86-88]. Structure-based virtual screening has been used for many years and the number of successful cases is increasing. Some successful examples are listed in Table 1.1.

Table 1.1 Some successful cases in virtual screening

Target	Algorithm	program	Reference
HIV-1 integrase	Lamarckian GA	Autodock	[89]
integrin $\alpha_v\beta_3$	Shape matching	Dock	[90]
G protein	Incremental construction	FlexX	[91]
Retinoic acid receptor	Monte-Carlo Minimization	ICM	[92]
Signal transducer and activator of transcription	Descriptor matching	Glide	[93]
Farnesyltransferase	Shape matching	EUDOCK	[94]

1.6.3.2 Ligand-based virtual screening

Ligand-based virtual screening identifies drug candidates by evaluating the similarity between compounds from chemical databases and known active compounds. This method is based on the principle that chemical structure is related to its biological property and is especially useful when there is no known 3D crystal structure of target proteins [95]. The methods of comparing structural similarity can be classified into three types: pharmacophore-based, molecular shape-based and molecular interaction field-based [96]. The pharmacophore method identifies compounds which can maintain the hydrophobic skeleton and hydrogen bond acceptors or donors in the reference ligand after docking with target protein [97]. Molecular shape-based method measures the overlap of molecules. In the field-based method, molecular interaction fields calculated from protein structure is used as the standard [98]. Ligand-based virtual screening can quickly identify lead compounds by directly screening the chemical databases.

1.7 Saturation-transfer difference nuclear magnetic resonance

Saturation-transfer difference nuclear magnetic resonance (STD-NMR) is a simple but powerful ligand-based technique to study the interaction between small molecules and protein. As shown in Figure 1.10, when the ^{1}H signal of a receptor is selectively irradiated by a pulse (Usually, the irradiated position is located at 0 to -1.0 ppm due to the appearance of resonances of methyl groups in folded proteins [99].) which does not excite the ligand, the saturation can be transferred to the ligand which binds to this receptor by spin diffusion through the nuclear overhauser effect (NOE) [100-101]. Namely, at first, the protein is selectively saturated, and then the magnetization of spin will recover to equilibrium via relax-action process. So the principle of NOE is when the ligands are close enough to proteins, they will also be involved in this process by intermolecular dipolar coupling and ^{1}H-^{1}H cross-relaxation [99, 102]. Then, the active ligands are dissociated back to solution and in the same time, more unsaturated lignads enter the binding pocket to gain saturation. The signal of protons in the binding fragment will be recorded in STD-NMR. On the contrast, there is no signal for the unbound ligand. This technique can confirm whether interaction exists between the small molecule and the receptor and which fragment on the small molecule is closer to the receptor, since the closest proton will gain the highest degree of saturation. Competitive STD-NMR can be used to probe whether ligands share the same bind site, as the STD-NMR signal of low-affinity binding ligand will reduce or disappear when a

competing high-affinity ligand is introduced [103]. Besides, it may also be applied to the measurement of binding constant between small molecule and receptor [103-104]. In this thesis, STD-NMR and competitive STD-NMR will be used to confirm the interaction between small molecule and protein and to test whether small molecules share the same binding site with MoeA.

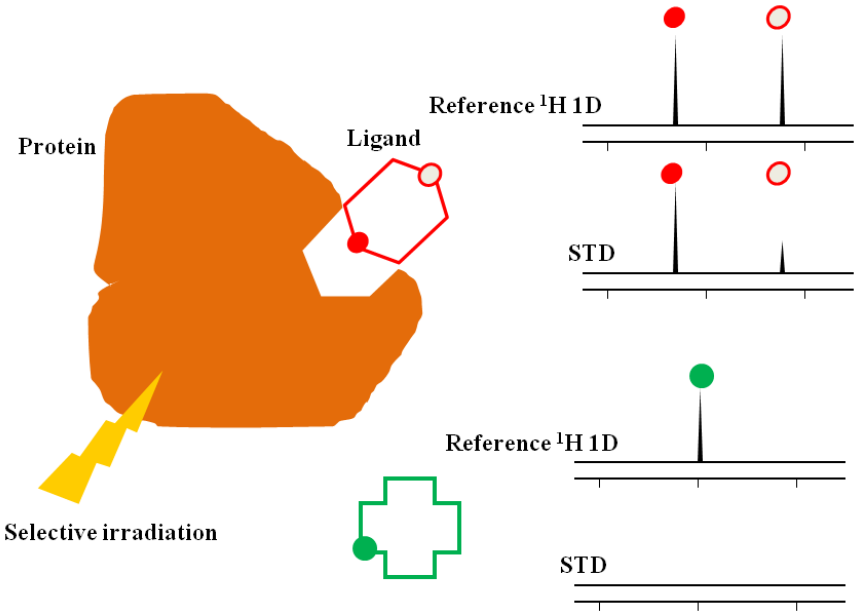


Figure 1.10 Principal of STD-NMR. Adopted from reference [105]. Signals of protons on ligand are recorded on STD-NMR spectra. There is no signal for unbound small molecule.

1.8 Aims and objectives

There is an urgent need to discover new antibacterial drugs based on new targets. Glycosyltransferase (GT) plays a significant role in the biosynthesis of bacteria cell wall and is a promising target for antibacterial drugs due to its conserved encoded gene and the absence of eukaryotic counterpart. However, no clinical used drug based on this target is known until today. The aim of this thesis is to discover potential small molecule antibacterial agents by virtual screening and improve their antibacterial activity through chemical modification of their structures, and then validate the interaction between these compounds and GT by biological assays.

The identification of hit compounds by virtual screening of chemical libraries against GT will be reported in chapter 2. The synthesis of a hit compound **GT10** and its analogues will be presented in chapter 3. Finally, the structure activity relationship will be discussed and the results of STD-NMR to testify the interaction between the optimized compound and the GT domain of *S. aureus* PBP 2 will be reported in chapter 4 of this thesis.

Chapter 2

Identification of hit compounds by virtual high throughput screening and bioassays

2.1 Introduction

High throughput screening (HTS) of large chemical libraries against biological targets is an essential step to identify active compounds in drug discovery. Although the conventional screening technique has been used in the discovery of some potential glycosyltransferase inhibitors, it is time-consuming and costly [106-108]. Molecular docking can simulate the binding process of receptor and ligand and identifies potential drug candidates on the basis of estimated binding affinities. Compared to conventional HTS, virtual screening based on molecular docking is fast and inexpensive. Furthermore, chemical synthesis of all compounds in the database is not necessary. The hit rate of docking-based virtual screening has been reported to be dozens of times higher than conventional HTS [109].

As mentioned in section 1.6.3, virtual screening can be either structure-based or ligand-based. Since moenomycin is the only known inhibitor that surely binds to the GT domain, it is very difficult to build a precise pharmcophore model for ligand-based virtual screening. On the contrary, the high resolution X-ray crystal structure of the GT domain of *Escherichia coli* PBP 1b, *Staphylococcus aureus* PBP 2 and *Aquifex aeolicus* PBP 1a have been obtained [32-34]. Thus, structure-based virtual screening is logically applied in the discovery of GT inhibitors.

Internal Coordinate Mechanics (ICM) is a docking software using Monte Carlo simulation that is widely used in drug discovery [110]. One advantage of ICM docking is that both receptor and ligand can be made flexible in the docking procedure. It provides several large-scale random moves, such as pseudo-Brownian moves, optimally biased moves of torsions and single torsion moves [111-112]. ICM now has been successfully applied in lead compound discovery [113]. In this chapter, 3,000,000 chemical structures were screened against the GT domain of *Staphylococcus aureus* by ICM pro. The top-ranking compounds were considered as hit compounds. Their antibacterial ability were then tested against several bacterial strains by measuring the minimum inhibitory concentration (MIC). The interaction between the active compound and the GT domain was confirmed by STD-NMR.

2.2 Experimental

2.2.1 Materials

2.2.1.1 Chemicals

Sodium phosphate monobasic dihydrate, kanamycin sulfate, EDTA, sodium hydroxide, Tris-HCl, deuterium oxide and sodium dodecyl sulfate were purchased from Sigma-Aldrich Co. Sodium chloride and isopropyl β-D-1-thiogalactopyranoside (IPTG) were obtained from USB Corporation. Reagents for Bradford protein assay were purchased from Bio-Rad Laboratories. Dimethyl sulfoxide-d_6 (DMSO-d_6) was obtained from Cambridge Isotope Laboratories. All the analytical grade solvents were purchased from Oriental Chemicals & Laboratory Supplies Ltd. and were used without further purification.

2.2.1.2 Media

Nutrient agar, tryptone and yeast extract were purchased from Oxoid Ltd. (Nepean, Ontario, Canada). 2xTY medium for protein expression was prepared by adding 2 g yeast, 3.2 g tryptone and 1.0 g sodium chloride into 200 mL deionized water and sterilized. Luria-Bertani (LB) medium was prepared by mixing 0.5 g yeast, 1.0 g tryptone, 0.5 g sodium chloride in 100 mL deionized water and sterilized. Müller-Hinton broth (MHB), cation-adjusted Müller-Hinton broth (CA-MHB) and trypticase soy broth (TSB) for minimum inhibitory concentration (MIC) test were purchased from Becton, Dickinson and Company (New Jersey, USA).

2.2.1.3 Bacterial strains and biochemicals

The *S. aureus* GT domain was expressed in *E. coli* BL21 (DE3) as a host. *S. aureus* ATCC 29213 was purchased from American Type Culture Collection (USA). *E. coli* ATCC 25922 and *B. subtilis* 168 are our laboratory collections. These three types of bacteria were used in the antibacterial test. Bovine serum albumin (BSA) and lysozyme were purchased from Sigma-Aldrich Co.

2.2.2 Virtual screening based on molecular docking

The GT domain of PBP consists of two sub-sites: a donor site for the elongating glycan chain and an acceptor site for the lipid II substrate. This domain can accommodate six sugar units. Between the sub-sites is a flexible region which plays an important role in the translocation of the product from the acceptor site to the donor site through a folding and unfolding process [114]. Several conserved residues, such as the glutamate residues Glu114 and Glu171 in *S. aureus* PBP 2, are essential for the catalysis of lipid polymerization of glycan chain [115]. Inhibition of the GT reaction can be achieved either by a compound binding to the active sites of the enzyme, like moenomycin which occupies the donor site mimicking the elongating glycan chain or by an agent binding to the lipid II substrate [33, 116]. The virtual ligand screening approach was used to identify a novel scaffold of small molecule binding to the GT catalytic binding site.

2.2.2.1 Protein preparation

All computational works were performed using ICM software version 3.6-1c [117]. The 2.8 Å X-ray crystal structure of GT-binding site of *S. aureus* PBP 2 in complex with moenomycin (PBD entry 2OLV) was retrieved from the RCSB Protein Data Bank (http://www.rcsb.org) and was prepared for the virtual ligand screening using the protein preparation protocol implemented in ICM.

MMFF (Merck Molecular Force Field) atom types were assigned to the protein atoms, cocrystal ligand, and water molecules [72]. Formal and partial charges were calculated at pH 7.4, and hydrogen atoms were added and optimized to the best hydrogen bonding network. The three possible protonation states of histidines, as well as 180° flips of histidine, glutamine, and aspargine residues were attempted to optimize the hydrogen bonds.

2.2.2.2 Ligand preparation

For ligand binding mode prediction, the cocrystal ligand moenomycin was extracted from the GT crystal structure of *S. aureus* PBP 2. The extracted ligand was converted into 1D using SMILES notation and then converted into 3D again and minimized using ICM software. This procedure generated starting conformations for docking that are independent from the cocrystal geometry. For virtual ligand screening, about 3,000,000 compounds from commercially available databases were downloaded using MolCart of ICM

software and were converted into 3D. Duplicate molecules were excluded. The compounds were filtered to remove molecules with known reactive groups and unfavorable absorption/permeation properties using ICM software. The Lipinski rule of five was used to remove compounds with molecular weight above 500, calculated octanol−water partition coefficient above 5, more than 5 hydrogen bond donors and more than 10 hydrogen bond acceptors [118]. Compounds with polar surface area above 120 $Å^2$ were also excluded. Compounds with Tanimoto distance (chemical dissimilarity) less than 0.3, defined as 1 − T, were excluded. For the remaining molecules, stereoisomers were generated for racemic compounds, and protonation states were calculated at pH 7.4 using the automatic pKa model implemented in ICM software.

2.2.2.3 Redocking of the known inhibitor moenomycin A

To validate the accuracy of ICM docking software for the subsequent virtual ligand screening, cocrystal ligand (moenomycin) was docked against the crystal structure of *S. aureus* GT binding site (PDB entry 2OLV). To account for the ligand-induced effect upon moenomycin binding to the GT domain, 50 conformations of the GT-binding site of *S. aureus* PBP 2 were generated. Briefly, induced-fit movements upon binding of moenomycin were simulated in the 50 conformations of GT-binding site using the ICM fully flexible docking algorithm [119]. Residues with side chain heavy atoms within a 5 Å cutoff distance from the cocrystal ligand were allowed to

randomly move using the ICM Biased Probability Monte Carlo algorithm, followed by a full local energy minimization [112].

The predicted binding mode of cocrystal ligand (moenomycin) was compared to the experimental coordinates using symmetry corrected RMSD that accounts for topological symmetries of chemical groups (e.g. equivalent carbons in benzene rings or equivalent negatively charged carboxylic oxygens) and three-dimensional symmetries generated by rotation. The conformation of GT binding site of *S. aureus* PBP 2 with the lowest RMSD assigned to the native ligand (moenomycin) was saved and subsequently used for virtual ligand screening of commercially available databases.

2.2.2.4 Docking-based virtual ligand screening.

Compounds of commercially available databases were docked against the best-performing conformation of GT binding site of *S. aureus* PBP 2 using the ICM software. The GT binding site was defined as the cavity delimited by residues with at least one non-hydrogen atom within a 4.0 Å cutoff radius from the co-crystal ligand moenomycin. In order to reduce the time of calculation, binding site residues were considered to be rigid (rigid-protein docking) and the pocket was represented by 0.5 Å grid potential maps accounting for hydrogen bonding, hydrophobic, van der Waals and electrostatic interactions [120]. The molecules were flexibly docked against the rigid model of the GT binding site and scored according to the ICM

scoring function [117]. Ligands receiving docking score values below a threshold of -32 are considered as good binders. Three independent docking runs were performed and the top-scoring pose of each compound was selected for further analysis. In order to maximize the chemical diversity for biological screening, binary fingerprints were calculated for each compound, and the molecules were clustered using the unweighted pair group method with arithmetic mean (UPGMA) agglomerative hierarchical algorithm implemented in the ICM software. About one hundred representative top-scoring compounds within clusters with a Tanimoto distance of at least 0.3 were selected for visual inspection. The predicted binding poses were compared with the experimental X-ray co-crystal structure of moenomycin. Thirty-four compounds establishing characteristic moenomycin-like interactions with the GT binding site were selected. The molecules were purchased and tested in the antimicrobial susceptibility test.

2.2.3 Antibacterial test

The MIC value of the hit compounds were measured by the antimicrobial susceptibility test using the broth microdilution procedure in accordance to the Clinical and Laboratory Standards Institute (CLSI) guidelines. Cation-adjusted Mueller Hinton broth (CA-MHB) for *S. aureus* strain ATCC 29213 and Mueller Hinton broth (MHB) for antibiotic-susceptible strains *B. subtilis* strain 168 and *E. coli* strain ATCC 25922 were used in the assays. Cells in exponential phase of growth were diluted to approximately 5×10^5 c.f.u/mL. Stock solutions of each compound were prepared in DMSO. Then, serial dilutions of each compound in CA-MHB or MHB medium were added into a 96-well microplate. The final percentage of DMSO in the assay was 1 %. Control experiments were performed with 1 % DMSO instead of the compound solution. After being incubated at 37 ^{0}C for 18 hours, OD_{600} value of cells was measured on a microplate reader (Bio-Rad laboratory Ltd., UK) and the percentage of bacterial cell inhibition with respect to controls was calculated. The MIC value was defined as the lowest compound concentration at which the growth of bacteria was inhibited by $\geq$ 90 %. Three independent assays were performed for each test.

2.2.4 Preparation of recombinant *S. aureus* GT domain

2.2.4.1 Expression and purification of *S. aureus* GT domain

The gene encoded GT domain (PBP 2 residues 76-243) was amplified from the full length PBP 2 by PCR and then integrated into vector pET24d for protein expression. After transformation of recombinant plasmid into *E. coli* BL21(DE3) cells with a 6-Histidine tag attached at its N-terminus under the control of a T7 promoter, it was streaked on a nutrient agar plate containing kanamycin (50 μgmL^{-1}) and the agar plate was incubated at 37 ºC overnight. The bacteria were grown in 2×TY medium containing kanamycin (50 μg mL^{-1}). A single colony was then inoculated into 5 mL of LB medium, which was then incubated at 37 ^{o}C with shaking at 250 rpm for 16 h. The overnight culture was transferred into a fresh 2×TY medium in a dilution ratio of 1:100 and kanamycin (50 μg mL^{-1}) was then added, followed by incubation at 37 ^{o}C with shaking at 250 rpm. When the OD_{600} reached 0.8, protein expression was induced by adding 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After induction, the cell culture was further incubated at 37 ^{o}C with shaking at 250 rpm for another 4 h. Cells were harvested by centrifugation at 4 ^{o}C, 9000 rpm for 20 min and then resuspended in 20 mM sodium phosphate buffer containing 500 mM NaCl (pH 7.4) followed by sonication. Debris anchored GT domain were collected by centrifugation and extracted by 20 mL starting buffer containing 20 mM FOS-choline-14 overnight at 4 ^{0}C. Almost all the targeted protein was found in the supernatant. The GT domain protein was then purified by Fast Protein

Liquid Chromatography (FPLC). The supernatant obtained by centrifugation at 13000 rpm at 4 ^{0}C for 45 minutes was filtered and then loaded on a 5 mL HiTrap chelating column (Amersham-Pharmacia) after it had been totally coated with nickel (II) sulfate solution and equilibrated with starting buffer (0.5 M NaCl, 0.02 M NaH_2PO_4, 1 mM FOS-Choline-14; pH 7.4). Finally, the product was separated by elution buffer (0.5 M NaCl, 0.02 M NaH_2PO_4, 1 mM FOS-Choline-14, 0.5 M imidazole; pH 7.4) and analyzed by SDS-PAGE. The collected protein was dialyzed against 20 mM sodium phosphate buffer (pH 7.4) and stored at -80 ^{0}C.

2.2.4.2 Sodium dodecyl sulfate-polyacrylaminde gel electrophoresis (SDS-PAGE) analysis

The protein fractions collected in different tubes were examined by 12 % SDS-PAGE in a Mini-PROTEAN III dual slab cell (Bio-Rad Laboratories). Firstly, to 10 µl sample was added reducing agent β-mercaptoethanol and SDS, and then the solution was boiled for 8 minutes at 100 ^{0}C. Subsequently, 10 µl mixtures were loaded on the SDS-PAGE gel made of 5 % stacking gel (pH 6.8) and 12 % separating gel (pH 8.8). Then, the samples were separated by electrophoresis at 200 V for 1 hour in running buffer (25 mM Tris-HCl, 192 mM glycine, 3.5 mM SDS). Finally, the SDS-PAGE gel was stained with coomassie blue for 5 minutes, then destained in methanol: glacial acetic acid: DIH_2O = (1:1:8) until the protein bands turned clear.

2.2.4.3 Electrospray ionization mass spectrometry (ESI-MS)

ESI-MS was used to measure the molecular weight of protein. Initially, the *S. aureus* GT domain was buffer exchanged into 20 mM ammonium acetate solution through a 15 mL Amicon Ultra filter centrifuge tube (10 KDa, Millipore Corporation). The protein mass spectra were recorded on a VG Platform quadrupole-time of flight (Q-TOF2) mass spectrometer (Micromass, Altrincham, Cheshire, UK). Protein sample was dissolved in H_2O/CH_3CN (1:1) with 1 % formic acid. After sample loading, the mass spectrum at the m/z 570-1600 range was recorded. The spectrometer was calibrated by myoglobin (10 ρmol/μl, average mass 16,951.5).

2.2.4.4 Determination of protein concentration

The concentration of protein was measured according to Bradford protein assay [121]. To an 800 µl protein sample was added 200 µl protein assay dye reagent (Bio-Rad Laboratories) and then the mixture was incubated at room temperature for 8 minutes. A solution of 800 µl DIH_2O and 200 µl protein assay dye reagent was used as background. After equilibrium, the absorbance was measured at 595 nm by a UV-Vis spectrophotometer. Ultimately, the protein concentration would be calculated from the standard curve calibrated by BSA.

2.2.5 Saturation-transfer difference (STD)-NMR

The STD-NMR experiments were performed on a Bruker AvanceIII 600 instrument equipped with a 5-mm QCI cryoprobe. Samples were prepared in 20 mM sodium phosphate in D_2O (pH 7.4) containing 5 % DMSO-d_6. The final concentration of GT domain and small molecule were 5 µM and 500 µM, respectively. Acquisitions were performed at 298 K using the standard STD pulse sequence with a train of 50 ms Gauss-shaped pulses, each separated by a 1 ms delay for selective protein irradiation, and an alternation between on and off resonances. A T1ρ spinlock filter (50 ms) was incorporated to suppress protein resonances. Presaturation of the protein resonances was performed with an on-resonance irradiation at -0.4 ppm; off resonance irradiation was applied at -30 ppm where no NMR resonances of ligand or protein are present. The STD signals were obtained by performing experiments at 5 s saturation and recorded with a recovery delay of 5 s to avoid incomplete relaxation to thermal equilibrium. Experiments were recorded using a minimum of 128 scans and 16 K data points. On- and off-resonance spectra were processed independently, and subtracted to provide a difference spectrum. Data was treated by TopSpin program suite (Bruker BioSpin Pte Ltd).

2.3 Results and discussion

2.3.1 Structure-based virtual screening of the potential GT inhibitors

2.3.1.1 Redocking of the known inhibitor moenomycin A

To validate the accuracy of docking algorithm for the subsequent virtual ligand screening process, re-docking of co-crystallized ligand moenomycin (Figure 2.1A) in the GT binding site of *S. aureus* PBP 2 (PDB entry 2OLV) was performed [33]. Due to the large ligand size and high number of rotatable bonds, 50 conformations of GT binding site were generated to re-dock moenomycin. The root mean square deviation (RMSD) between the experimental binding pose and the predicted binding pose of co-crystal ligand moenomycin in each conformation was calculated. This value accounts for topological symmetries of chemical groups (e.g. equivalent carbons in phenyl rings or equivalent negatively charged carboxylic oxygens) and three dimensional symmetries generated by rotation. Values below 2.5 Å are considered acceptable. A lower value of RMSD means a good docking. The lowest RMSD value of 50 conformations was 0.58 Å, which is much better than the minimum value of 2.5 Å. The ICM binding score of moenomycin A was −34.71. Superimposition of the docked conformation and the crystallographic conformation of moenomycin indicated good overlap between the two conformations (Figure 2.1B). The predicted binding pose of moenomycin in the GT binding site of *S. aureus* PBP 2 is shown in Figure 2.1C.

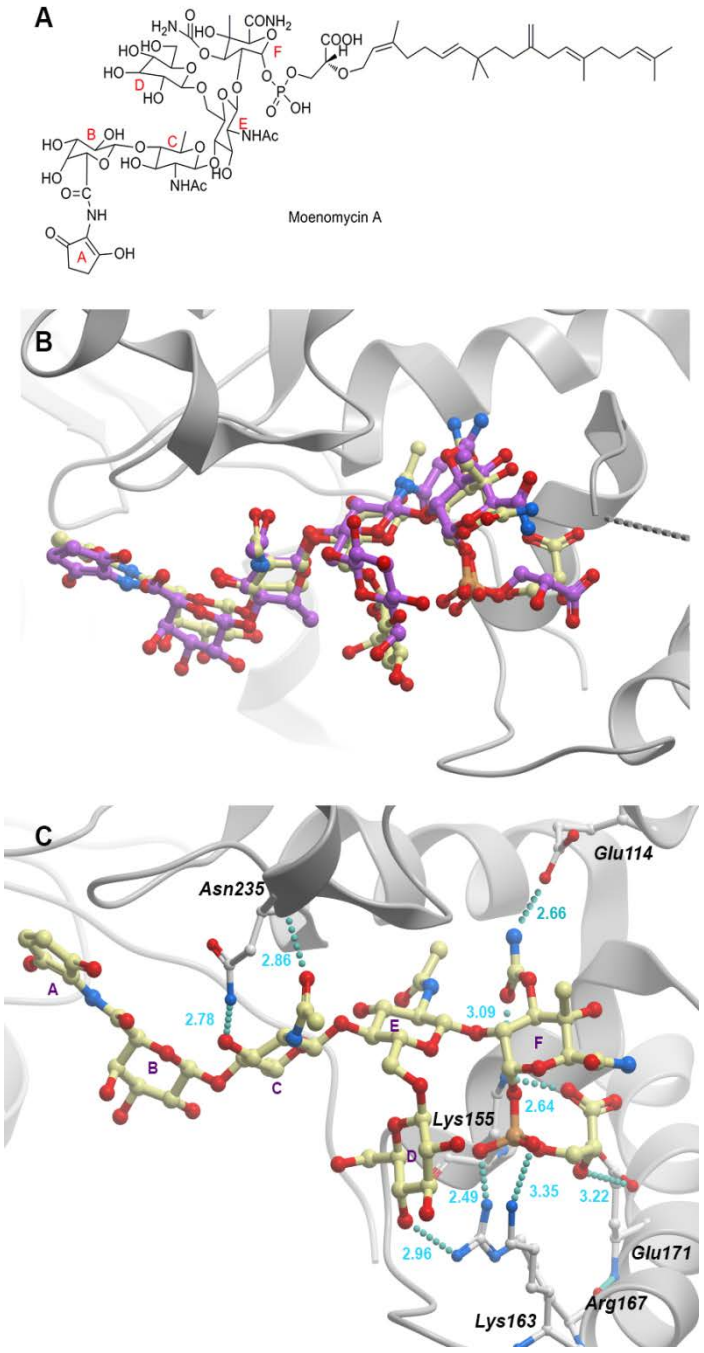


Figure 2.1 (A) Chemical structure of moenomycin. (B) Overlay of predicted binding pose and crystallographic binding pose of moenomycin. The prediction is represented with yellow carbons and red oxygens, whereas the co-crystal ligand is represented with purple carbons and red oxygens. (C) Predicted binding pose of moenomycin to the GT binding site of *S. aureus* PBP 2 (PDB entry 2OLV). Hydrogen bonds are represented with cyan spheres. The labeled interacting residues are shown with a "ball and stick" model and colored by atom types.

Consistent with the X-ray crystallographic data [33], moenomycin is predicted to bind in the cleft between the head and jaw regions of the GT binding site through the formation of hydrogen bonds with the critical residues Glu114 (2.66 Å), Lys155 (2.64 Å and 3.09 Å), Lys163 (3.35 Å), Arg167 (2.96 Å and 2.49 Å), Glu171 (3.22 Å) and Asn235 (2.78 Å and 2.86 Å). The best-performing conformation of GT binding site was then used for subsequent virtual ligand screening of approximately 3 million compounds from commercially available databases.

2.3.1.2 *In silico* screening and compounds selection

We conducted large scale molecular docking-based library screening to identify compounds that could bind to the GT binding site of *S. aureus* PBP 2. A flexible-ligand/rigid receptor docking in ICM software was used to screen approximately 3,000,000 compounds from the commercially available databases. The binding pocket of the protein (receptor) is represented by a set of grid maps, and each compound (ligand) of the database was docked to the pocket. A score reflecting the “fitting quality” of ligand-receptor complex was calculated. After the virtual ligand screening, the best scoring compounds were inspected visually and evaluated according to their chemical and drug-like properties as well as three-dimensional conformations of the docked ligand-receptor complex. The thirty-four top-scoring compounds were selected and purchased from vendors. The

structures and ICM binding scores of selected compounds are shown in Figure 2.2.

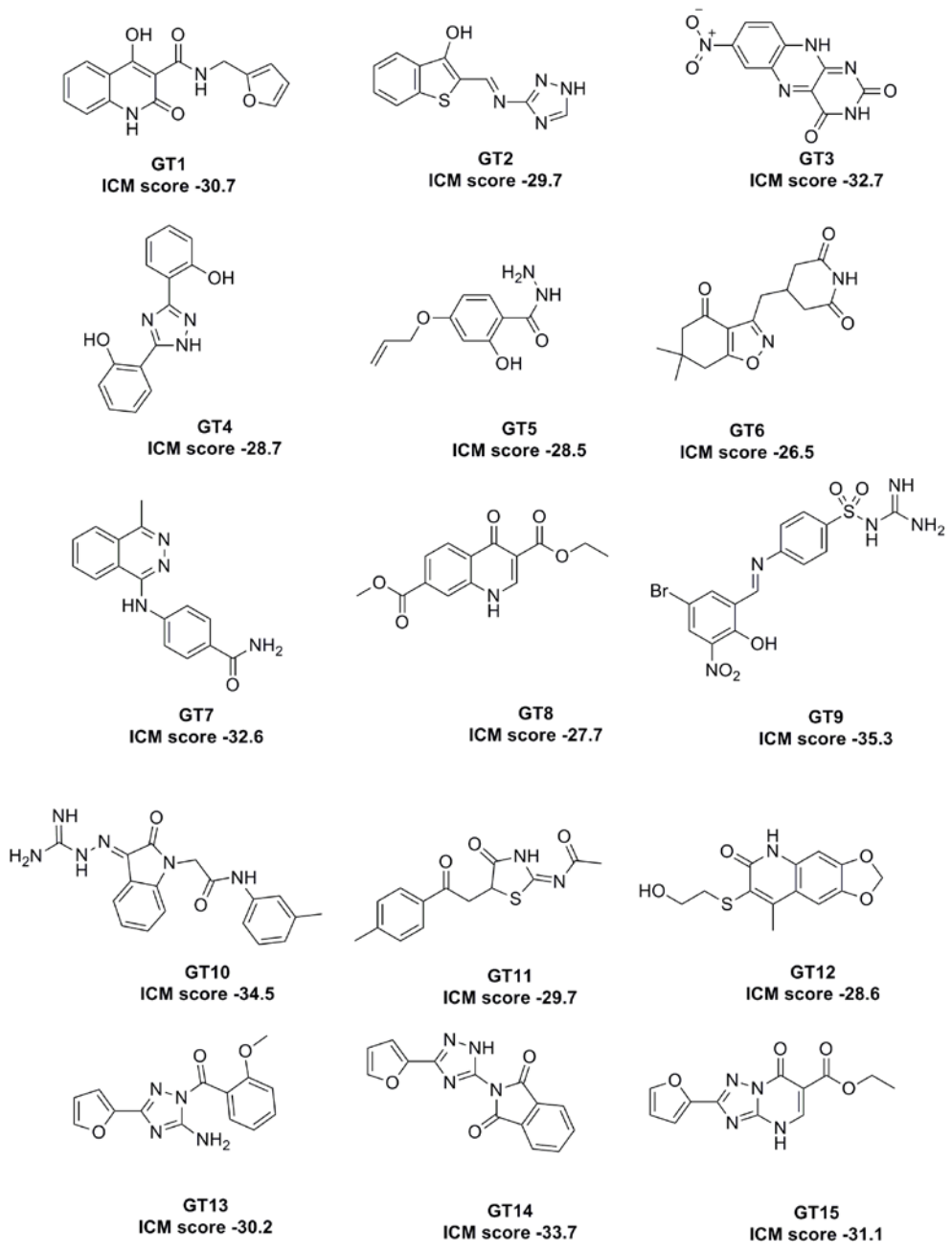
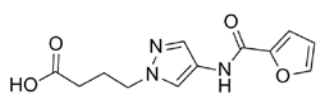
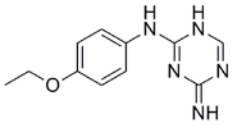


Figure 2.2 Structures and ICM binding scores of the virtual screening hit compounds **GT1-GT34**.

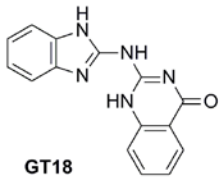
Continued…



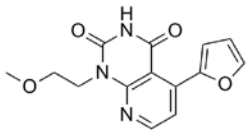
GT16
ICM score -28.5



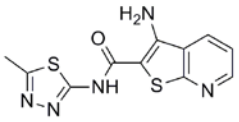
GT17
ICM score -29.3



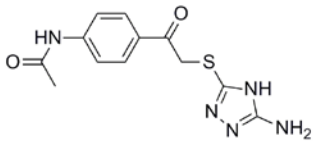
GT18
ICM score -27.2



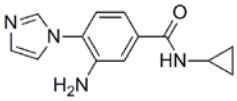
GT19
ICM score -29.9



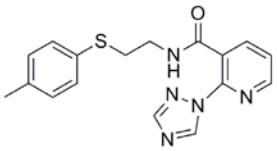
GT20
ICM score -31.2



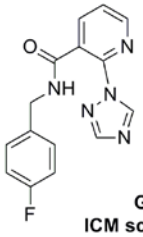
GT21
ICM score -31.9



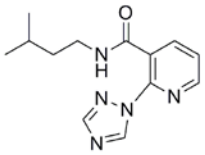
GT22
ICM score -26.9



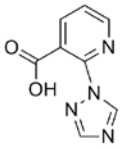
GT23
ICM score -28.1



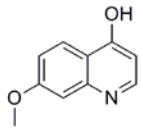
GT24
ICM score -26.7



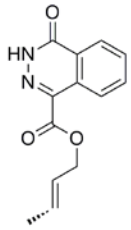
GT25
ICM score -25.4



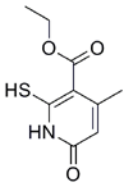
GT26
ICM score -24.4



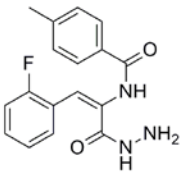
GT27
ICM score -23.1



GT28
ICM score -26.2

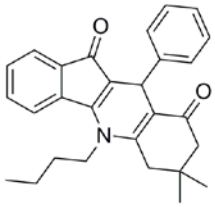


GT29
ICM score -27.9

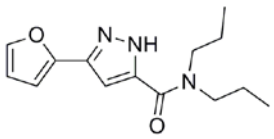


GT30
ICM score -32.1

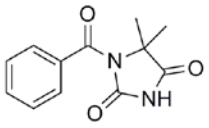
Continued…



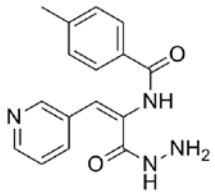
GT31
ICM score -26.9



GT32
ICM score -27.1



GT33
ICM score -25.9



GT34
ICM score -30.9

2.3.2 Antibacterial activity of hit compounds obtained from virtual screening

The antibacterial ability of the 34 hit compounds obtained by virtual screening were tested against Gram-positive strains *S. aureus*, *B. subtilis* and Gram-negative strain *E. coli*. The results are summarized in Table 2.1. It was noted that compound **GT10** could be used for further modification as it showed moderate antibacterial activity (192 µg mL^{-1} against *S. aureus* and *B. subtilis*). The predicted binding pose of **GT10** is shown in Figure 2.3A. Docking analysis showed that three hydrogen bonds are formed between **GT10** and active site residues, including the nitrogen atoms of guanidyl moiety of **GT10** interacting with the main chain oxygen atom of Lys155 (3.01 Å) and the main chain oxygen atom of Ser160 (2.89 Å and 3.18 Å). The conserved positively charged Lys155 residue located near the donor pyrophosphate region of the GT active site was proposed to translocate a newly formed glycan chain to the donor site [33]. In addition, a *meta*-methylbenzyl group of **GT10** establishes hydrophobic interactions with the side chains of Ile195, Tyr196 and Val233. Hydrophobic interaction between the isatin core of **GT10** and the side chain of residue Ile195 was also observed. As shown in Figure 2.3B, **GT10** partly overlaps with co-crystallized moenomycin rings C and E, but is buried deeper via formation of hydrophobic interactions with residues Ile195, Tyr196 and Val233. In contrast, moenomycin does not form hydrophobic interaction in the hydrophobic shelf. Besides forming strong hydrogen and hydrophobic

interactions, the molecular geometry of **GT10** favors its fit into the GT binding pocket of *S. aureus* PBP 2 that occupies the region of growing glycan chain on the basis of the proposed mechanism for lipid II polymerization [33]. Those interactions may be responsible for the antibacterial activity of **GT10**.

Table 2.1 MICs of hit compounds obtained from virtual screening

Compound	MIC, μg mL^{-1}		
	S. aureus 29213	*B. subtilis 168*	*E. coli 25922*
GT1	>384	>384	>384
GT2	>384	>384	>384
GT3	>384	>384	>384
GT4	>384	>384	>384
GT5	>384	>384	>384
GT6	>384	>384	>384
GT7	>384	>384	>384
GT8	>384	>384	>384
GT9	384	384	>384
GT10	192	192	>384
GT11	>384	>384	>384
GT12	>384	>384	>384
GT13	>384	>384	>384
GT14	>384	>384	>384
GT15	>384	>384	>384
GT16	>384	>384	>384
GT17	>384	>384	>384
GT18	>384	>384	>384

Continued…

Compound	MIC, µg mL^{-1}		
	S. aureus 29213	*B. subtilis 168*	*E. coli 25922*
GT19	>384	>384	>384
GT20	>384	>384	>384
GT21	>384	>384	>384
GT22	>384	>384	>384
GT23	>384	>384	>384
GT24	>384	>384	>384
GT25	>384	>384	>384
GT26	>384	>384	>384
GT27	>384	>384	>384
GT28	>384	>384	>384
GT29	>384	>384	>384
GT30	>384	>384	>384
GT31	>384	>384	>384
GT32	>384	>384	>384
GT33	>384	>384	>384
GT34	>384	>384	>384

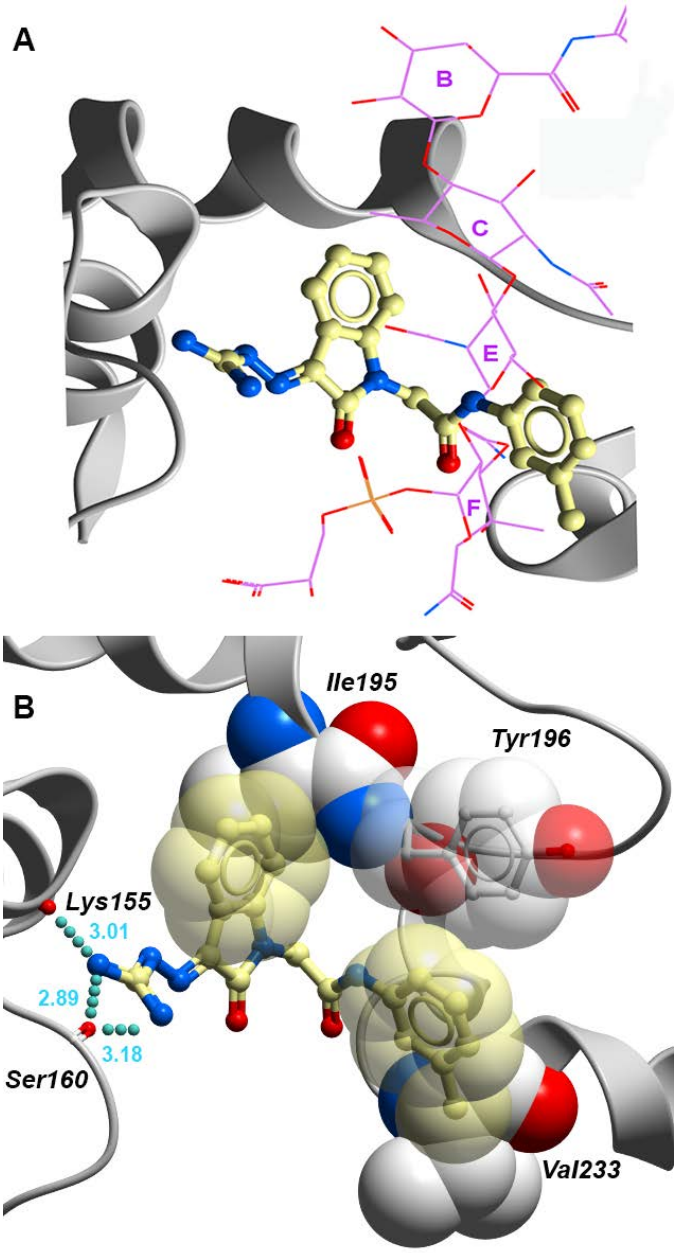


Figure 2.3 (A) Superimposition of the predicted binding pose of compound **GT10** (yellow carbons) to the crystal structure of moenomycin (purple wire, meonomycin rings are labeled from B to F) bound to *S. aureus* PBP 2 (PDB entry 2OLV). (B) The predicted binding pose of compound **GT10** to the GT binding site of *S. aureus* PBP 2. Hydrophobic residues and hydrophobic moieties of compound **GT10** are shown in space-fillling form. Hydrogen bonds are represented with cyan spheres. The labeled interacting residues are shown with a "ball and stick" model and colored by atom types.

2.3.3 Expression of *S. aureus* GT domain

The *S. aureus* GT domain was successfully expressed with a good yield according to the procedures in section 2.2.4.1. The chromatogram and 12 % SDS-PAGE results of the elution fractions obtained by FPLC are shown in Figure 2.4 and Figure 2.5 respectively. The SDS-PAGE result indicated that the purity of the target protein is acceptable (above 80 %). Finally, the experimental molecular mass of *S. aureus* GT domain is consistent with the theoretical value (20020.39 Da). (Figure 2.6)

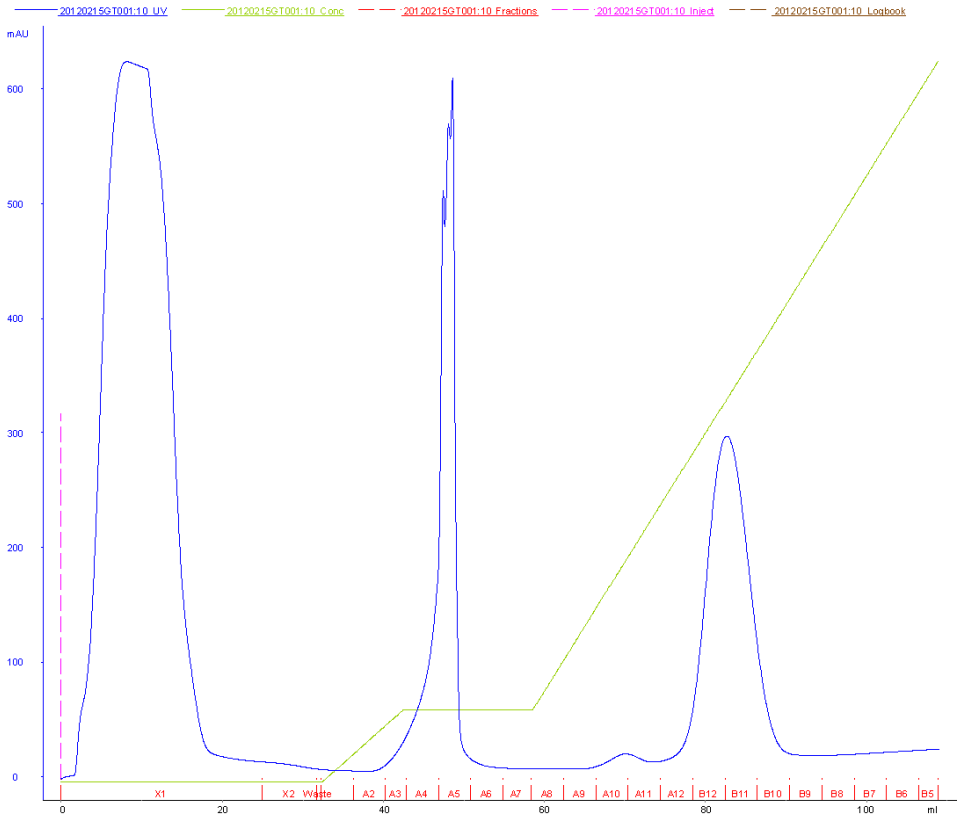


Figure 2.4 FPLC chromatogram of *S.aureus* GT domain.

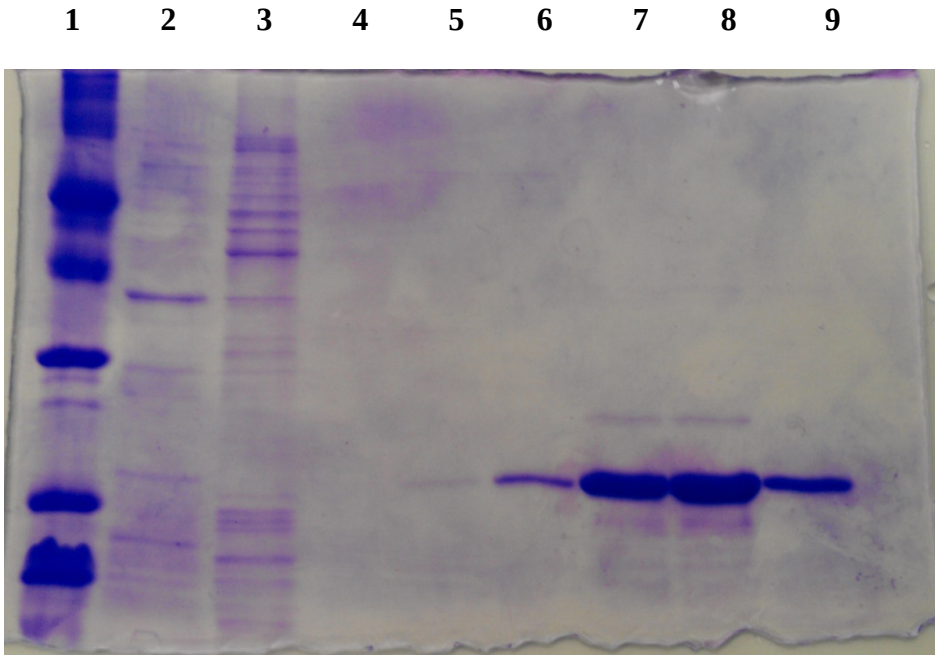


Figure 2.5 Electrophoresis results of 12 % SDS-PAGE. Lane 1, low range molecular marker: rabbit muscle phosphorylase b (97400 Da), BSA (66200 Da), hen egg white ovalbumin (45000 Da), bovine carbonic anhydrolase (31000 Da), soybean trypsin inhibitor (21500 Da), hen egg white lysozyme (14400 Da); Lane 2, flow through; Lane 3, non-specific binding proteins; Lane 4-5, elution in A_{10}, A_{11}; Lane 6-9, elution of *S. aureus* His-tagged GT domain.

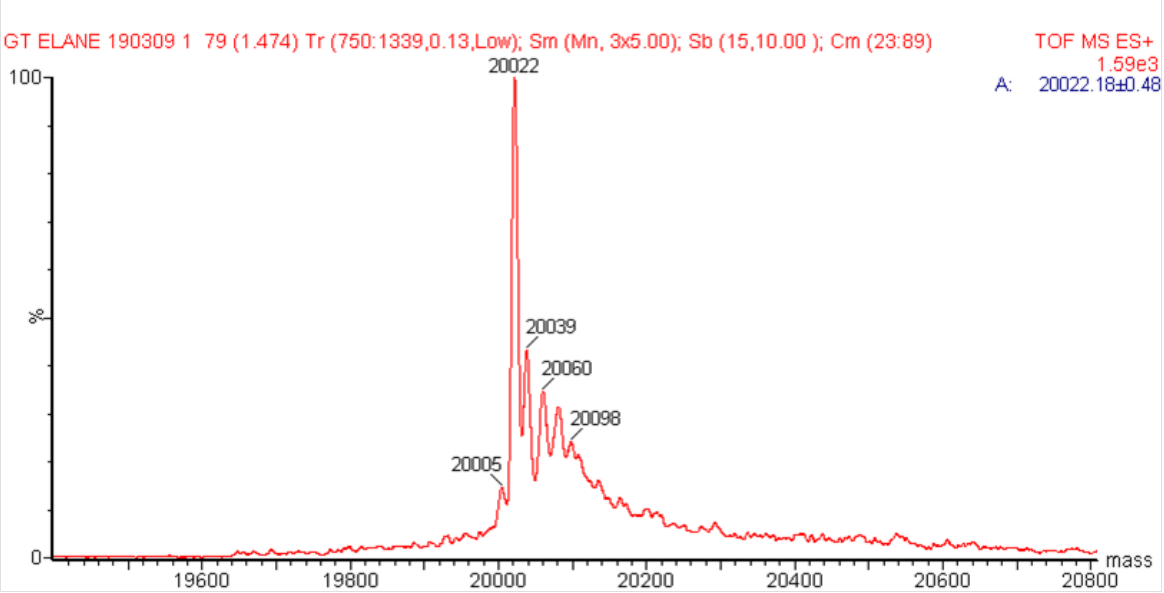
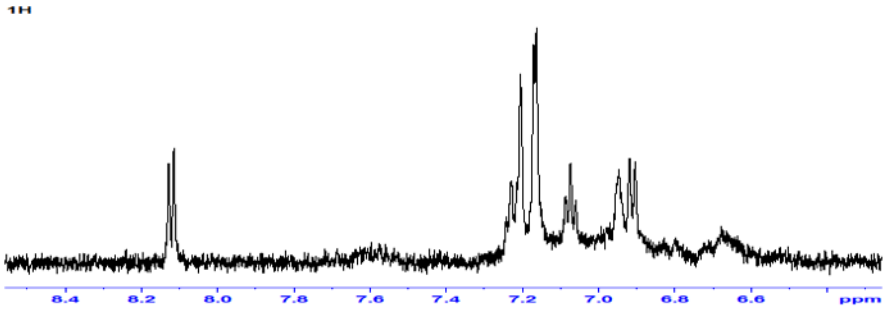


Figure 2.6 ESI-mass spectrum of *S. aureus* His-tagged GT domain.

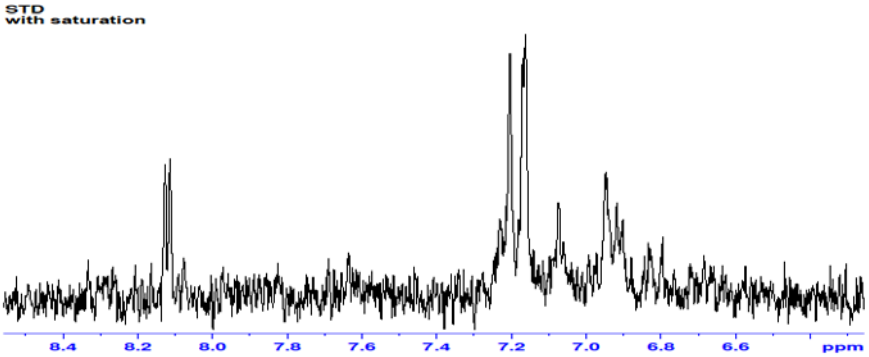
2.3.4 Saturation-transfer difference (STD)-NMR

STD-NMR is a very useful tool to detect the interaction between small molecule and protein. When the protein is selectively irradiated to saturation, this saturation will be transferred to the protons which are in close contact with the protein on the small molecule. The results indicated that when the GT domain was saturated (-0.4 ppm), this saturation could be transferred to the protons on the phenyl group in **GT10 (**Figure 2.7B). On the other hand, when the irradiation was performed at -30 ppm where no protein or ligand signal was present, the STD signal disappeared (Figure 2.7C). The off resonance experiment indirectly proved the binding between **GT10** and the GT domain. STD-NMR showed that the two phenyl groups of **GT10** may enter the binding site of GT domain. However, the STD-NMR signal is relatively weak and hard to be correctly assigned due to its poor solubility in aqueous solution and weak binding affinity (Figure 2.7).

A



B



C

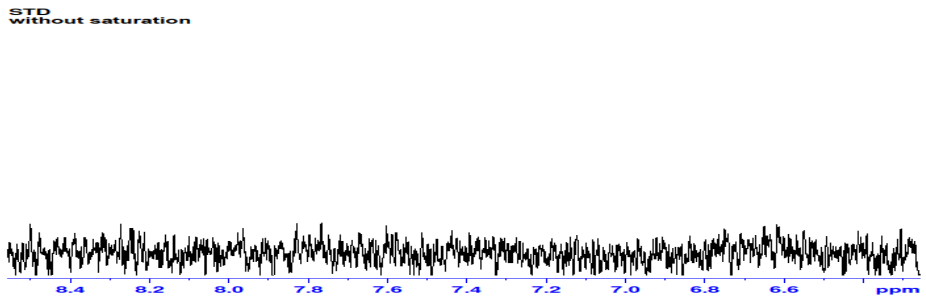


Figure 2.7 (A) ^{1}H-NMR spectrum of **GT10**; (B) STD-NMR spectrum between **GT10** and GT domain; (C) Off resonance STD-NMR spectrum.

2.4 Concluding remarks

Virtual screening is a very useful technique to assist the discovery of novel inhibitors. In this chapter, the best-performing conformation for predicting the binding pose of moenomycin was selected as the initial docking template for virtual ligand screening. Combination of virtual screening of commercially available compounds with *in vitro* biological assays identified compound **GT10** as a novel scaffold of GT inhibitor. The generated docking pose showed that **GT10** interacts with several amino acid residues that are critical for binding to the growing glycan chain in linear glycan polymerization. Good shape complementarity of **GT10** with the active site was also shown in our docking model. The identified hit compound **GT10** was experimentally tested for binding to the GT binding site of *S. aureus* PBP 2 by STD-NMR experiment. A scaffold of the hit compound **GT10** was then used for design and synthesis of its derivatives and the results are reported in the following chapter.

Chapter 3

Synthesis and characterization of GT10 and its analogues

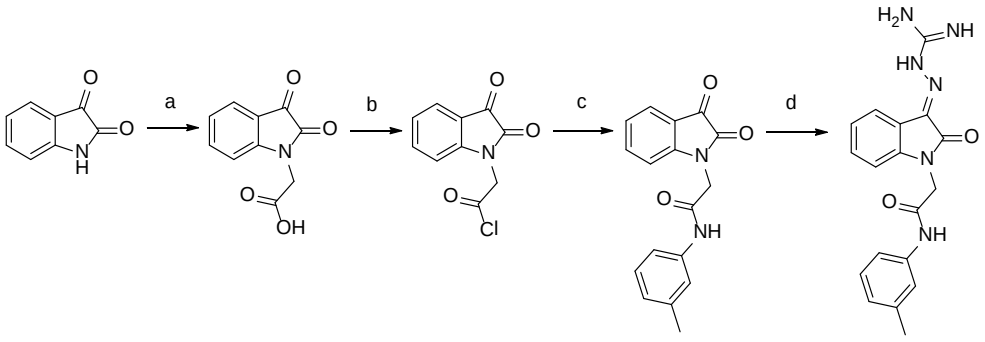
3.1 Introduction

The main structure of the hit compound **GT10** obtained from virtual screening is an N-substituted Schiff base of isatin. Isatin is a natural product from plants such as *Isatis tincotoria*, *Calanthe discolor* and *Couroupita guianensis* [122]. Besides, it is also found as an endogenous indole in the brain, peripheral tissues and body fluids in mammal [123]. Isatin is the basic structure of many natural products and synthetic compounds which exhibit a variety of biological effects such as antiviral, antifungal, antiproliferative, and even antibacterial activity [124-127]. Yet, the mechanisms of those activities are still unknown.

Therefore, the isatin structure in **GT10** is considered as the basic pharmacore. In order to optimize the structure of this hit compound and investigate the binding property of GT domain, a number of **GT10** analogues were designed with different side chains and functional groups (Table 3.1). Moreover, derivatives with replacement of guanidyl by phenyl groups or aliphatic ring were also designed. The antibacterial property against different bacterial strains of these analogues will provide valuable information on lead optimization based on **GT10**.

In this chapter, a general synthetic method for **GT10** analogues is reported (Figure 3.1). The compounds synthesized were characterized by ^{1}H-NMR and low resolution mass spectra (appendix I & II).

(A)



(B)

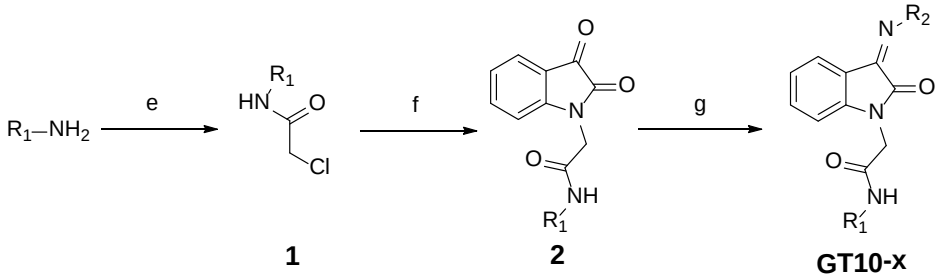
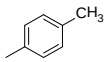
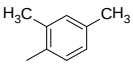
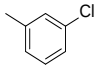
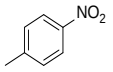
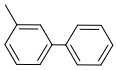
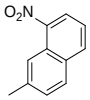


Figure 3.1 Synthetic routes for compound **GT10** and its analogues. **a)** Ethyl bromoacetate, NaOH; b**)** $SOCl_2$; **c)** m-Toluidine; **d)** Aminoguanidine hydrochloride; **e)** Chloroacetyl chloride, CH_3CN, 0^0C; **f)** Isatin, K_2CO_3, DMF, 80 ^0C ; **g)** R_2-NH_2, CH_3COOH, reflux or CH_3CH_2OH, reflux

Table 3.1 GT10 and its analogues

Compounds	R_1	R_2
GT10		
GT10-1		
GT10-2		
GT10-3		
GT10-4		
GT10-5		
GT10-6		
GT10-7		
GT10-8		
GT10-9		
GT10-10		
GT10-11		
GT10-12		
GT10-13		
GT10-14		
GT10-15		
GT10-16		
GT10-17		
GT10-18		
GT10-19		

Continued…

Compounds	R_1	R_2
GT10-20		
GT10-21		
GT10-22		
GT10-23		
GT10-24		
GT10-25		
GT10-26		
GT10-27		
GT10-28		
GT10-29		

3.2 Design of GT10 analogues

According to the results of STD-NMR, the two phenyl groups in **GT10** have interaction with the GT domain. Substances with isatin as basic structure possess diverse biological effects. Thus, the isatin core was kept as pharmacophore in the design of new **GT10** analogues. Firstly, compounds (**GT10-14** to **GT10-17)** were designed to testify whether the guanidyl group could be replaced. We also investigated the importance of phenyl groups in R_1 by substituting it with some aliphatic chains (**GT10-1** to **GT10-8**). As it was proven by STD-NMR that the phenyl group in R_1 binds to GT, we designed compounds with the same substitution but at different positions on the phenyl group (**GT10**, **GT10-10**, and **GT10-11)**. The trend of antibacterial ability of these three compounds was further testified by comparing with the antibacterial ability of **GT10-19**, **GT10-20** and **GT10-22.** After it had been confirmed that the phenyl group in R_1 was vital to biological effect and the substitutions should be restricted to the *meta*-position, a variety of analogues with different *meta*-substituted functional groups were synthesized to search for antibacterial agents with higher activity. Finally, attempts were made to increase the hydrophobic interaction between small molecules and GT domain by introducing biphenyl and substituted naphthyl group into R_1.

3.3 Experimental

3.3.1 Materials

Cyclopentylamine, cyclohexylamine, ethylamine, propylamine, allylamine, butylamine, furfruylamine, 2-thiophenemethylamine, *m*-toluidine, *o*-toluidine, *p*-toluidine, aniline, 2, 4-dimethylanline, benzylamine, *m*-fluoroaniline, *o*-fluoroaniline, *p*-fluoroaniline, 3-chloroaniline, 4-nitroaniline, 3-nitroaniline, 3,4-(methylenedioxy)aniline, 3-picolylamine, 3-aminobenzonitrile, *m*-anisidine, 3-aminobiphenyl and 2-amino-1-naphthalenesulfonic acid were purchased from Sigma-Aldrich Co. 8-Nitro-2-naphthylamine was synthesized by using 2-amino-1-naphthalenesulfonic acid as starting material [128]. The deuterated solvents were purchased from Cambridge Isotope Laboratories, Inc. and Meryer (Shanghai) Chemical Technology Co. Ltd. All solvents were purchased from Oriental Chemicals & Lab. Supplies Ltd. Silicon gel (70-200 μm) was used as stationary phase in column chromatography. All the chemicals and solvents mentioned above were directly used without any further purification.

3.3.2 Instrumentation

^{1}H-NMR spectra were recorded on a Bruker Avancelll 400 MHz spectrometer in CD_3OD with the residual solvent peak as a reference standard. Chemical shifts (δ) and coupling constants (*J*) are expressed in parts per million. All low resolution mass spectra were measured in HPLC grade methanol by a Waters Acquity ESI mass spectrometry with Single Quad (SQ) detector.

3.3.3 General procedures for the synthesis of GT10 analogues

To mixture of amines (1.0 mmol) and potassium carbonate (1.2 mmol) in CH_3CN at 0 ^0C, chloroacetyl chloride (1.1 mmol) was added by several portions. After being kept at 0 ^0C for an additional 2 hours, the solution was poured into water and then extracted by CH_3CN. Compound **1** was obtained by removal of solvent and used for the next step without further purification.

Subsequently, potassium carbonate (0.6 mmol) was added to the solution of isatin (0.5 mmol) in DMF in an ice bath, and the resulting mixture was kept at room temperature for another 30 minutes. Then compound **1** (0.5 mmol) in DMF was added. Before poured into diluted HCl solution, the mixture was heated at 80 ^0C for 4~6 hours and monitored by TLC with a mixture of ethyl acetate and petroleum ether as developing solvent. The precipitated product was collected by filtration and purified by column chromatography using a mixture of ethyl acetate and petroleum ether as eluent.

Finally, compound **2** was reacted with the corresponding amines in acetic acid or ethanol under reflux to yield compounds **GT10-x**, which were purified by column chromatography with a mixture of CH_3OH and CH_2Cl_2 as eluent or by recrystallization from ethanol.

3.4 Results and discussion

3.4.1 Product yields and characterization

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(m-tolyl) acetamide (GT10). Yield 76 %. ^{1}H NMR (400 MHz, CD_3OD) δ: 2.33(s, 3H), 4.69(s, 2H), 6.94-6.97(m, 2H), 7.09(t, *J*=8.0Hz, 1H), 7.20(t, *J*=7.6Hz, 1H), 7.27(t, *J*=7.6Hz, 1H), 7.36(d, *J*=8.4, 1H), 7.41(s, 1H), 8.38(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{18}H_{19}N_6O_2$ m/z: 351.16 $[M+H]^+$; Found 350.76

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-cyclopentyl acetamide (GT10-1). Yield 60 %. ^{1}H NMR (400 MHz, CD_3OD) δ: 1.48-1.55(m, 2H), 1.62-1.68(m, 2H), 1.71-1.79(m, 2H), 1.92-2.00(m, 2H), 4.12-4.19(m, 1H), 4.49(s, 2H), 6.97(d, *J*=8.0Hz, 1H), 7.21(t, *J*=7.6Hz, 1H), 7.49(t, *J*=7.6Hz, 1H), 7.80(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{16}H_{21}N_6O_2$ m/z: 329.17 $[M+H]^+$; Found 328.96

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-cyclohexyl acetamide (GT10-2). Yield 74 %. ^{1}H NMR (400MHz, CD_3OD) δ: 1.18-1.42(m, 5H), 1.66(d, broad, 1H), 1.78(d, broad, 2H), 1.88(d, broad, 2H), 3.67-3.73(m, 1H), 4.49(s, 2H), 6.97(d, *J*=8.0Hz, 1H), 7.21(t, *J*=7.6Hz, 1H), 7.49(t, *J*=8.0Hz, 1H), 7.81(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{17}H_{23}N_6O_2$ m/z: 343.19 $[M+H]^+$; Found 342.99

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-ethylacetamide (GT10-3). Yield 55 %. ^{1}H NMR (400MHz, CD_3OD) δ: 1.15(t, *J*=7.2Hz, 3H), 3.24-3.33(m, 2H), 4.50(s, 2H), 6.99(d, *J*=8.0Hz, 1H), 7.21(t, *J*=7.6Hz, 1H), 7.49(t, *J*=7.6Hz, 1H), 7.81(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{13}H_{17}N_6O_2$ m/z: 289.14 $[M+H]^+$; Found 388.91

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-propyl acetamide (GT10-4). Yield 64 %. ^{1}H NMR (400 MHz, CD_3OD) δ: 0.94(t, *J*=7.2Hz, 3H), 1.51-1.60(m, 2H), 3.20(t, *J*=6.8Hz, 2H), 4.51(s, 2H), 6.99(d, *J*=8.0Hz, 1H), 7.21(t, *J*=7.6Hz, 1H), 7.48(t, *J*=8.0Hz, 1H), 7.80(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{14}H_{19}N_6O_2$ m/z: 303.16 $[M+H]^+$; Found 302.96

N-allyl-2-(3-(2-carbamimidoylhydrazono)-2-oxoindolin-1-yl)acetamide (GT10-5). Yield 48 %. ^{1}H NMR (400M Hz, CD_3OD) δ: 3.86(d, *J*=4.8Hz, 2H), 4.55(s, 2H), 5.13(d, *J*=10Hz, 1H), 5.22(d, *J*=16.8Hz, 1H), 5.81-5.91(m, 1H), 7.01(d, *J*=8.0Hz, 1H), 7.22(t, *J*=7.6Hz, 1H), 7.50(t, *J*=8.0Hz, 1H), 7.81(d, *J*=7.2Hz, 1H); ESI-MS Calcd for $C_{14}H_{17}N_6O_2$ m/z: 301.14 $[M+H]^+$; Found 300.95

N-butyl-2-(3-(2-carbamimidoylhydrazono)-2-oxoindolin-1-yl)acetamide (GT10-6). Yield 57 %. ^{1}H NMR (400M Hz, CD_3OD) δ: 0.95(t, *J*=7.2Hz, 3H), 1.32-1.41(m, 2H), 1.48-1.56(m, 2H), 3.24(t, *J*=6.8Hz, 2H), 4.51(s, 2H), 6.98(d, *J*=8.0Hz, 1H), 7.21(t, *J*=7.6Hz, 1H), 7.48(t, *J*=7.6Hz, 1H), 7.80(d, *J*=7.2Hz, 1H); ESI-MS Calcd for $C_{15}H_{21}N_6O_2$ m/z: 317.17 $[M+H]^+$; Found 316.96

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(furan-2-ylmethyl) acetamide (GT10-7). Yield 43 %. ^{1}H NMR (400M Hz, CD_3OD) δ: 4.42(s, 2H), 4.54(s, 2H), 6.27(s, 1H), 6.36(s, 1H), 6.97(d, *J*=8.0Hz, 1H), 7.21(t, *J*=7.6Hz, 1H), 7.45-7.49(m, 2H), 7.80(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{16}H_{17}N_6O_3$ m/z: 341.14 $[M+H]^+$; Found 340.93

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(thiophen-2-ylmethyl) acetamide (GT10-8). Yield 55 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.54(s, 2H), 4.60(s, 2H), 6.94-7.01(m, 3H), 7.21(t, *J*=7.6Hz, 1H), 7.31(d, *J*=5.2Hz, 1H), 7.46(t, *J*=7.6Hz, 1H), 7.80(d, *J*=7.2Hz, 1H); ESI-MS Calcd for $C_{16}H_{17}N_6O_2S$ m/z: 357.11 $[M+H]^+$; Found 356.93

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-phenyl acetamide (GT10-9). Yield 75 %. ^{1}H NMR (400M Hz, CD_3OD) δ: 4.71(s, 2H), 7.09(d, *J*=8.0Hz, 1H), 7.14(t, *J*=7.6Hz, 1H), 7.22(t, *J*=7.6Hz, 1H), 7.34(t, *J*=7.6Hz, 2H), 7.50(t, *J*=8.0Hz, 1H), 7.57(d, *J*=8.0Hz, 2H), 7.82(d, *J*=7.6Hz, 1H) ; ESI-MS Calcd for $C_{17}H_{17}N_6O_2$ m/z: 337.14 $[M+H]^+$; Found 336.96

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(o-tolyl) acetamide (GT10-10). Yield 70 %. ^{1}H NMR (400MHz, CD_3OD) δ: 2.28(s, 3H), 4.76(s, 2H), 7.13(d, *J*=8.0Hz, 1H), 7.17-7.28(m, 4H), 7.34(d, *J*=8.8Hz, 1H), 7.53(t, *J*=8.0Hz, 1H), 7.82(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{18}H_{19}N_6O_2$ m/z: 351.16 $[M+H]^+$; Found 350.94

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(p-tolyl) acetamide (GT10-11). Yield 69 %. ^{1}H NMR (400MHz, CD_3OD) δ: 2.32(s, 3H), 4.69(s, 2H), 7.08(d, *J*=8.0Hz, 1H), 7.15(d, *J*=8.4Hz, 2H), 7.22(t, *J*=7.2Hz, 1H), 7.44(d, *J*=8.0Hz, 2H), 7.50(t, *J*=8.0Hz, 1H), 7.82(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{18}H_{19}N_6O_2$ m/z: 351.16 $[M+H]^+$; Found 350.96

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(2,4-dimethyl phenyl)acetamide (GT10-12). Yield 71 %. ^{1}H NMR (400MHz, CD_3OD) δ: 2.23(s, 3H), 2.30(s, 3H), 4.74(s, 2H), 7.01(d, *J*=8.4Hz, 1H), 7.08(s, 1H), 7.12(d, *J*=8.0Hz, 1H), 7.18(d, *J*=8.0Hz, 1H), 7.23(t, *J*=8.0Hz, 1H), 7.52(t, *J*=7.2Hz, 1H), 7.82(d, *J*=7.2Hz, 1H); ESI-MS Calcd for $C_{19}H_{21}N_6O_2$ m/z: 365.17 $[M+H]^+$; Found 364.95

N-benzyl-2-(3-(2-carbamimidoylhydrazono)-2-oxoindolin-1-yl) acetamide (GT10-13). Yield 61 %. ^{1}H NMR (400M Hz, CD_3OD) δ: 4.43(s, 2H), 4.57(s, 2H), 6.99(d, *J*=8.0Hz, 1H), 7.22(t, *J*=7.6Hz, 1H), 7.27-7.35(m, 5H), 7.48(t, *J*=7.6Hz, 1H), 7.80(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{18}H_{19}N_6O_2$ m/z: 351.16 $[M+H]^+$; Found 350.96

2-(2-Oxo-3-(phenylimino)indolin-1-yl)-N-(m-tolyl)acetamide (GT10-14). Yield 54 %. ^{1}H NMR (400MHz, CD_3OD) δ: 2.34(s, 3H), 4.71(s, 2H), 6.65(d, *J*=7.6Hz, 1H), 6.83(t, *J*=7.6Hz, 1H), 6.97(d, *J*=7.6Hz, 1H), 7.02(d, *J*=8.8Hz, 1H), 7.06(d, *J*=8.0Hz, 2H), 7.22(t, *J*=8.0Hz, 1H), 7.31-7.46(m, 4H), 7.52(t, *J*=7.2Hz, 2H); ESI-MS Calcd for $C_{23}H_{19}N_3NaO_2$ m/z: 392.14 $[M+Na]^+$; Found 391.91

2-(3-((3-Fluorophenyl)imino)-2-oxoindolin-1-yl)-N-(m-tolyl)acetamide (GT10-15). Yield 56 %. ^{1}H NMR (400MHz, CD_3OD) δ: 2.34(s, 3H), 4.71(s, 2H), 6.68(d, *J*=7.6Hz, 1H), 6.84-6.90(m, 3H), 6.97(d, *J*=7.2Hz, 1H), 7.04-7.09(m, 2H), 7.21(t, *J*=7.6Hz, 1H), 7.37(d, *J*=7.6Hz, 1H), 7.41(s, 1H), 7.46(t, *J*=8.0Hz, 1H), 7.50-7.56(m, 1H); ESI-MS Calcd for $C_{23}H_{18}FN_3NaO_2$ m/z: 410.13 $[M+Na]^+$; Found 409.91

2-(3-(Cyclohexylimino)-2-oxoindolin-1-yl)-N-(m-tolyl)acetamide (GT10-16). Yield 40 %. ^{1}H NMR (400MHz, CD_3OD) δ: 1.29-1.39(m, 1H), 1.44-1.66(m, 4H), 1.75(d, broad, 1H), 1.84-1.89(m, 4H), 2.33(s, 3H), 4.61(s, 2H), 5.21-5.26(m, 1H), 6.94-6.97(m, 2H), 7.15(t, *J*=7.6Hz, 1H), 7.20(t, *J*=7.6Hz, 1H), 7.35(d, *J*=8.4Hz,1H), 7.39(s, 1H), 7.46(d, *J*=8.0Hz,1H), 7.72(d, *J*=7.2Hz, 1H); ESI-MS Calcd for $C_{23}H_{25}N_3NaO_2$ m/z: 398.18 $[M+Na]^+$; Found 397.97

2-(3-((3-Chlorophenyl)imino)-2-oxoindolin-1-yl)-N-(m-tolyl)acetamide (GT10-17). Yield 53 %. ^{1}H NMR (400MHz, CD_3OD) δ: 2.34(s, 3H), 4.71(s, 2H), 6.67(d, *J*=7.6Hz, 1H), 6.88(t, *J*=7.6Hz, 1H), 6.97(d, *J*=6.8Hz, 1H), 7.02(d, *J*=7.6Hz, 1H), 7.06(d, *J*=7.6Hz, 1H), 7.13(s, 1H), 7.22(t, *J*=7.6Hz, 1H), 7.24-7.36(m, 2H), 7.38(s, 1H), 7.41-7.52(m, 2H); ESI-MS Calcd for $C_{23}H_{17}ClN_3O_2$ m/z: 402.10 $[M-H]^-$; Found 401.92

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(4-nitrophenyl) acetamide (GT10-18). Yield 40 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.77(s, 2H), 7.10(d, *J*=8.0Hz, 1H), 7.23(t, *J*=7.6Hz, 1H), 7.50(t, *J*=7.6Hz, 1H), 7.82-7.86(m, 3H), 8.25(d, *J*=9.2Hz, 2H); ESI-MS Calcd for $C_{17}H_{16}N_7O_4$ m/z: 382.13 $[M+H]^+$; Found 381.92

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(4-fluoro phenyl)acetamide (GT10-19). Yield 71 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.70(s, 2H), 7.06-7.11(m, 3H), 7.22(t, *J*=7.6Hz, 1H), 7.48-7.59(m, 3H), 7.82(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{17}H_{16}FN_6O_2$ m/z: 355.13 $[M+H]^+$; Found 354.91

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(2-fluoro phenyl)acetamide (GT10-20). Yield 65 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.78(s, 2H), 7.09(d, *J*=8.0Hz, 1H), 7.15-7.24(m, 4H), 7.51(t, *J*=8.0Hz, 1H), 7.82(d, *J*=7.6Hz, 1H), 7.88(t, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{17}H_{16}FN_6O_2$ m/z: 355.13 $[M+H]^+$; Found 354.93

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(3-chloro phenyl)acetamide (GT10-21). Yield 56 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.72(s, 2H), 7.09(d, *J*=8.0Hz, 1H), 7.14(d, *J*=9.2Hz, 1H), 7.23(t, *J*=8.0Hz, 1H), 7.32(t, *J*=8.0Hz, 1H), 7.45(d, *J*=8.4Hz, 1H), 7.50(t, *J*=7.6Hz, 1H), 7.74(s, 1H), 7.83(d, *J*=7.2Hz, 1H); ESI-MS Calcd for $C_{17}H_{16}N_6O_2Cl$ m/z: 371.10 $[M+H]^+$; Found 370.87

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(3-fluoro phenyl)acetamide (GT10-22). Yield 64 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.72(s, 2H), 6.87(t, *J*=7.6Hz, 1H), 7.08(d, *J*=8.0Hz, 1H), 7.23(t, *J*=7.2Hz, 1H), 7.28-7.38(m, 2H), 7.48-7.54(m, 2H), 7.82(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{17}H_{16}FN_6O_2$ m/z: 355.13 $[M+H]^+$; Found 354.92

N-(benzo[*d*][1,3]dioxol-5-yl)-2-(3-(2-carbamimidoylhydrazono)-2-oxoindolin-1-yl)acetamide (GT10-23). Yield 55 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.67(s, 2H), 5.95(s, 2H), 6.79(d, *J*=8.4Hz, 1H), 6.94(d, *J*=8.4Hz, 1H), 7.07(d, *J*=8.0Hz, 1H), 7.21-7.24(m, 2H), 7.50(t, *J*=7.6Hz, 1H), 7.82(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{18}H_{17}N_6O_4$ m/z: 381.13 $[M+H]^+$; Found 380.95

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(pyridin-3-ylmethyl)acetamide (GT10-24). Yield 53 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.55(s, 2H), 4.64(s, 2H), 7.07(d, *J*=7.6Hz, 1H), 7.21(t, *J*=7.6Hz, 1H), 7.33(t, *J*=6.8Hz, 1H), 7.40(d, *J*=8.0Hz, 1H), 7.49(t, *J*=7.6Hz, 1H), 7.80-7.84(m, 2H), 8.51(s, 1H); ESI-MS Calcd for $C_{17}H_{18}N_7O_2$ m/z: 352.15 $[M+H]^+$; Found 351.94

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(3-cyano phenyl) acetamide (GT10-25). Yield 56 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.74(s, 2H), 7. 01(d, *J*=8.0Hz, 1H), 7.23(t, *J*=7.6Hz, 1H), 7.48-7.56(m, 3H), 7.81-7.84(t, *J*=7.2Hz, 2H), 8.05(s, 1H); ESI-MS Calcd for $C_{18}H_{16}N_7O_2$ m/z: 362.14 $[M+H]^+$; Found 361.94

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(3-methoxy phenyl)acetamide (GT10-26). Yield 66 %. ^{1}H NMR (400MHz, CD_3OD) δ: 3.79(s, 3H), 4.70(s, 2H), 6.72(d, *J*=8.0Hz, 1H), 7.08(d, *J*=8.0Hz, 1H), 7.20-7.28(m, 3H), 7.50(t, *J*=7.6Hz, 1H), 7.82(d, *J*=7.6Hz, 1H); ESI-MS Calcd for $C_{18}H_{19}N_6O_3$ m/z: 367.15 $[M+H]^+$; Found 366.95

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(3-nitrophenyl) acetamide (GT10-27). Yield 57 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.77(s, 2H), 7.11(d, *J*=8.0Hz, 1H), 7.23(t, *J*=7.6Hz, 1H), 7.50(t, *J*=7.6Hz, 1H), 7.60(t, *J*=8.0Hz, 1H), 7.83(d, *J*=7.6Hz, 1H), 7.93(d, *J*=6.8Hz, 1H), 8.00(d, *J*=8.0Hz, 1H), 8.62(s, 1H); ESI-MS Calcd for $C_{17}H_{16}N_7O_4$ m/z: 382.13 $[M+H]^+$; Found 381.92

N-([1,1'-biphenyl]-3-yl)-2-(3-(2-carbamimidoylhydrazono)-2-oxoindolin -1-yl)acetamide (GT10-28). Yield 56 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.75(s, 2H), 7.11(d, *J*=7.6Hz, 1H), 7.23(t, *J*=7.6Hz, 1H), 7.36(t, *J*=7.2Hz, 1H), 7.41-7.46(m, 4H), 7.51(t, *J*=6.8Hz, 1H), 7.56(d, *J*=6.4Hz, 1H), 7.61(d, *J*=7.2Hz, 2H), 7.83(d, *J*=7.2Hz, 1H), 7.86(s, 1H); ESI-MS Calcd for $C_{23}H_{21}N_6O_2$ m/z: 413.17 $[M+H]^+$; Found 412.96

2-(3-(2-Carbamimidoylhydrazono)-2-oxoindolin-1-yl)-N-(8-nitro naphthalene-2-yl)acetamide (GT10-29). Yield 50 %. ^{1}H NMR (400MHz, CD_3OD) δ: 4.80(s, 2H), 7.13(d, *J*=8.0Hz, 1H), 7.23(t, *J*=7.6Hz, 1H), 7.51(t, *J*=8.0Hz, 1H), 7.57(t, *J*=8.0Hz, 1H), 7.83(d, *J*=7.6Hz, 1H), 7.95(d, *J*=7.6Hz, 1H), 8.06(d, *J*=9.2Hz, 1H), 8.21(d, *J*=8.4Hz, 1H), 8.28(d, *J*=7.6Hz, 1H), 8.84(s, 1H); ESI-MS Calcd for $C_{21}H_{18}N_7O_4$ m/z: 432.14 $[M+H]^+$; Found 431.97

3.4.2 Synthetic routes for GT10 and its analogues

In general, amide bond can be formed by reacting amine with acyl chloride which can be generated from carboxylic acid. The synthesis of **GT10** could be successfully completed by such strategy as shown in route A of Figure 3.1 [129]. However, in the synthesis of **GT10-x** the intermediate acyl chloride generated by such procedure was difficult to handle due to hydrolysis, and the yields were low (approximately 20 %) under a harsh reaction condition. Thus, we revised the synthetic route on the basis of reported methods (route B in Figure 3.1) [130-131]. Compound **1** was obtained by reacting the corresponding amine with excess chloroacetyl chloride which is cheap and commercially available. Compound **1** was then connected to isatin to generate compound **2** by heating in DMF. Finally, the targeted compound **GT10-x** could be produced by condensation between compound **2** and the relevant amines (Figure 3.1). Through this new synthetic route, all the designed compounds were successfully synthesized with acceptable yields (40-76 %). Compared to the original synthetic method, the revised one is easier to handle and has higher yields due to a decrease in synthetic steps.

The synthesis of compound **1** went smoothly with very high yield. The relevant amine easily reacted with 10 % excess chloroacetyl chloride that could be easily removed by water. The impurities produced in this step were easy to be eliminated and would not affect the next step of the reaction, so the product was directly used in the next step after it had been dried under

vacuum. TLC showed only one spot under 254 nm UV light. However, it was difficult to obtain product when the reaction was conducted with an *ortho*-substituted electro-withdrawing functional group, such as the nitro group on R_1.

In the second step, bromoacetyl chloride could be used to replace chloroacetyl chloride due to its higher activity but the latter is easier to handle. Potassium iodide was added to accelerate the reaction speed and improve product yield. Since iodide is a good nucleophile, it was used to substitute chloride in compound **1**, and the generated product became more active towards isatin as iodide is also a good leaving group. This reaction occurs even when pyridine is introduced in compound **1**, and under such situation water could be used to quench the reaction instead of hydrochloric acid to minimize product loss. Compound **2** could be purified by column chromatography or recrystallization from acetone.

Finally, condensation between compound **2** and aminoguanidine hydrochloride was easy to carry out in the presence of hot acetic acid to give the final product as a yellow or orange solid. However, with some amines such as cyclohexanamine, the product would decompose during purification by column chromatography. So recrystallization was used to purify those products.

3.5 Concluding remarks

In summary, a variety of **GT10** analogues were designed on the basis of compound **GT10**, the hit compound obtained from structure-based virtual screening against GT domain of *S. aureus* PBP 2. By applying the revised synthetic route, all these targeted compounds were successfully synthesized with acceptable yields. Their structures were validated by ^{1}H-NMR and low resolution mass spectrometry. The structure activity relationship of these new synthetic substances will be discussed in chapter 4.

Chapter 4

Structure activity relationship study and molecular docking simulation

4.1 Introduction

Structure activity relationship (SAR), the relation between the three dimensional structure of molecules and their biological activity, can reveal the properties of the biological receptor, which in turn guides the design and optimization of new inhibitors in drug discovery. Together with molecular modeling, SAR can provide a better understanding of the ligand-receptor interaction.

As mentioned in chapter 3, a series of **GT10** analogues have been synthesized. In this chapter, their antibacterial activity against *S. aureus*, *B. subtilis* and *E. coli* are reported and a structure-activity relationship is proposed. This SAR study of **GT10** derivatives led to the identification of a more active GT inhibitor, **GT10-27**, which was shown by STD-NMR to interact with the GT domain of *S. aureus* PBP 2 and share the same binding pocket with moenomycin A. Finally, a docking simulation of **GT10-27** with the *S. aureus* GT domain binding pocket was conducted to reveal the binding pose of the compound.

4.2 Experimental

Antibacterial test of the **GT10** analogues were conducted as mentioned in section 2.2.3.

STD-NMR measurement was conducted according to the procedures in section 2.2.5. The final concentrations of the GT domain and **GT10-27** in the STD-NMR experiment were 2.5 μm and 250 μm respectively. Competition studies were performed by adding 2.5 to 5 μM of moenomycin A to the GT domain-small molecule mixture and the STD-NMR spectra were acquired under the same condition.

In the molecular docking simulation experiment, **GT10-22** and **GT10-27**, the two most active **GT10** analogues, were docked against the best-performing conformation of GT binding site of *S. aureus* PBP 2 reported in section 2.3.1.1. The procedure of docking simulation has been described in section 2.2.2.4. Briefly, a flexible ligand (**GT10-22** or **GT10-27**) was placed to the GT binding site. The top-scoring pose of each compound was selected to undergo the refinement process of the receptor pocket around the ligand. The refinement procedure was based on the BPMC global energy optimization. During the entire refinement process, the temperature was set at 600 K. The side chain variables of residues within 5 Å from the ligand were sampled. In this refinement step, the backbone and side chain variables of receptor with the free variables of the ligand could be effectively optimized by Monte Carlo sampling. After the refinement process, the

energy of each conformation was re-assessed. The most energetically favourable ligand-receptor pairs were re-docked and rescored.

Cytotoxicity of the **GT10-27** was measured against the cell lines A375, HFF-1 and Detroit 551 obtained from ATCC. A375, HFF-1and Detroit 551 are human melanoma cell line, normal human foreskin fibroblast and normal human skin fibroblast respectively. A375 was maintained in DMEM containing 10 % fetal bovine serum (FBS), 100 U ml^{-1} penicillin and 100 µg mL^{-1} streptomycin. HFF-1 was cultured in DMEM containing 15 % FBS, 100 U ml^{-1} penicillin and 100 µg ml^{-1} streptomycin. Detroit 551 was maintained in EMEM containing 10 % FBS, 100 U ml^{-1} penicillin and 100 µg ml^{-1} streptomycin. All cell lines were cultured at 37 °C in a humidified atmosphere of 5 % CO_2. To examine the anti-proliferation effect of the antibiotics on mammalian cells, 5,000 cells/well were seeded on a 96-well plate. After 24 h incubation, the culture medium was replaced with fresh medium containing various concentrations of the antibiotics. Cells were further incubated for 72 h and then 5 µl of 10 mg mL^{-1} water soluble MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) was added to each well and incubated for a further 4 h at 37 °C. Viable cells would convert the water soluble MTT into an insoluble formazan, which was then solubilized by SDS-HCl solution. After 18 h incubation at 37 °C, the concentration was determined by optical density at 570 nm with background at 655 nm. Finally, IC_{50} was calculated from the dose-response curves using GraphPad Prism 5.0.

4.3 Results and discussion

4.3.1 Structure activity relationship study

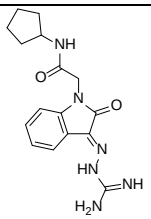
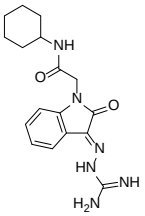
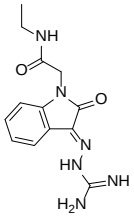
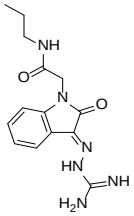
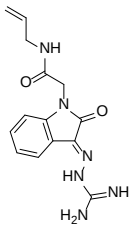
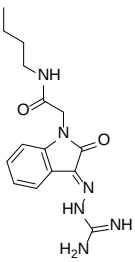
The MIC results of all the **GT10** analogues against Gram-positive strains (*S. aureus*, *B. subtilis*) and Gram-negative strain (*E. coli*) are summarized in Table 4.1. The MIC of the well-known GT inhibitor, moenomycin A, against *S. aureus*, *B. subtilis* and *E. coli*, are 0.12, 0.95 and 127 μg mL^{-1} respectively [60]. It is a strong inhibitor against Gram-positive bacteria, but rather weak against Gram-negative bacteria. As shown in Table 4.1, **GT10-27** shows the most potent antibacterial ability against *B. subtilis* with MIC at 24 μg mL^{-1} which is 8-fold higher than that of the parent compound **GT10**. It also inhibits the growth of *S. aureus* and *E. coli* with MIC values of 48 μg mL^{-1} and 96 μg mL^{-1} respectively. The antibacterial ability of **GT10-27** against *E. coli* is about 1.5-fold higher than that of moenomycin A.

When the nitro group in **GT10-27** was replaced by a fluorine group as in **GT10-22**, an ether group as in **GT10-23**, or a methoxyl group as in **GT10-26**, the antibacterial ability was reduced by over 50 %. Even worse was when the nitro group was substituted by a cyano group as in **GT10-25**, or by chlorine as in **GT10-21**, the antibacterial ability was lost with the MIC > 192 μg mL^{-1}. Furthermore, comparing the chemical structure of **GT10**, **GT10-10** and **GT10-11**, or **GT10-19**, **GT10-20** and **GT10-22**, the substitution on the phenyl group in R_1 should be restricted to *meta*-position to gain reasonable antibacterial ability. When the phenyl group R_1 was replaced by aliphatic chains as in **GT10-1** to **GT10-8**, a benzyl group as in

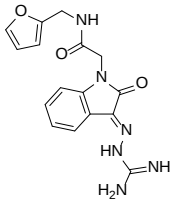
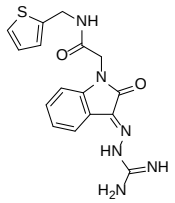
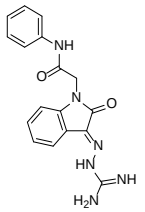
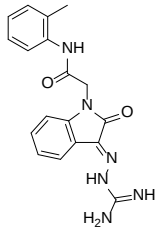
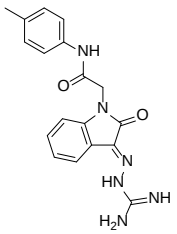
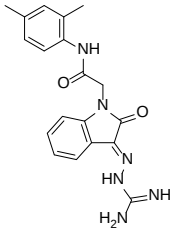
GT10-13 or pyridin-3-ylmethyl as in **GT10-24**, the antibacterial ability was lost against all strains probably because the rigid phenyl group in R_1 provides hydrophobic interaction with the GT binding site. Also, the unsubstituted **GT10-9** or disubstitued R_1 in **GT10-12** both led to a decrease of antibacterial ability. Besides, comparing the chemical structure of **GT10-27**, **GT10-28** and **GT10-29**, the biphenyl and naphthyl rings were too bulky to enter the GT binding site which led to a significant reduction in the antibacterial ability. The guanidyl group in R_2 is also essential. The antibacterial ability was reduced when the guanidyl group was replaced by cycloalkyl or phenyl groups (**GT10-14** to **GT10-17**). **GT10-22** shows a two-fold higher in antibacterial ability against *E. coli* (48 μg mL^{-1}) than the positive control moenomycin A (127 μg mL^{-1}) and **GT10-27** (96 μg mL^{-1}). This may be contributed to the fact that the fluorine group could facilitate the penetration of the compound through the cell membrane of Gram-negative strain.

In summary, **GT10-27** was found to possess strong antibacterial ability against Gram-positive and Gram-negative strains. The binding between this compound and the GT binding site was therefore further investigated.

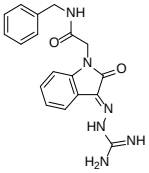
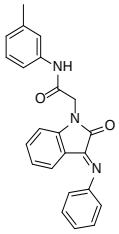
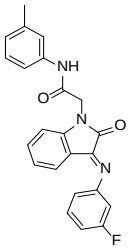
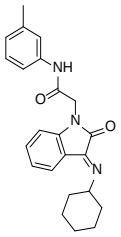
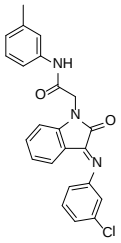
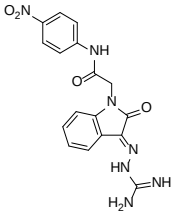
Table 4.1 MIC results of analogues of compound GT10

Compounds	Structures	MIC, μg mL^{-1}		
		S.aureus 29213	*B.subtilis 168*	*E.coli 25922*
GT10-1		>192	>192	>192
GT10-2		>128	>128	>128
GT10-3		>192	>192	>192
GT10-4		>192	>192	>192
GT10-5		>192	>192	>192
GT10-6		>192	>192	>192

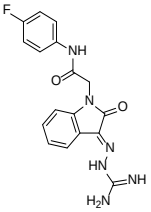
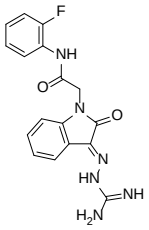
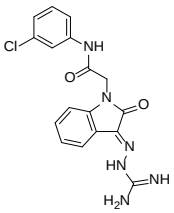
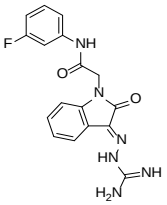
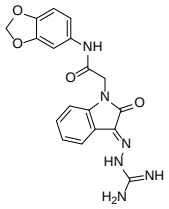
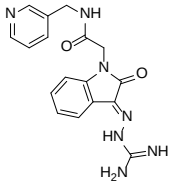
Continued…

Compounds	Structures	MIC, µg mL^{-1}		
		S.aureus 29213	*B.subtilis 168*	*E.coli 25922*
GT10-7		>192	>192	>192
GT10-8		>192	>192	>192
GT10-9		>128	>128	>128
GT10-10		>192	>192	>192
GT10-11		>192	>192	>192
GT10-12		>96	>96	>96

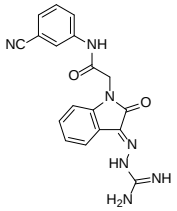
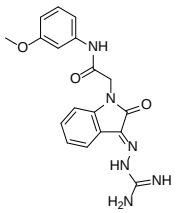
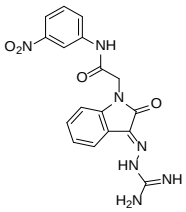
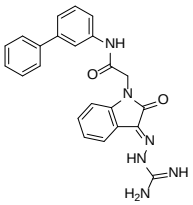
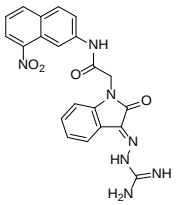
Continued…

Compounds	Structures	MIC, µg mL^{-1}		
		S.aureus 29213	*B.subtilis 168*	*E.coli 25922*
GT10-13		>192	>192	>192
GT10-14		>192	>192	>192
GT10-15		>192	>192	>192
GT10-16		96	>192	>192
GT10-17		>96	>96	>96
GT10-18		>96	>96	>96

Continued…

Compounds	Structures	MIC, µg mL^{-1}		
		S.aureus 29213	*B.subtilis 168*	*E.coli 25922*
GT10-19		96	48	96
GT10-20		>96	>96	>96
GT10-21		>192	>192	>192
GT10-22		96	48	48
GT10-23		192	96	192
GT10-24		>192	>192	>192

Continued…

Compounds	Structures	MIC, µg mL^{-1}		
		S.aureus 29213	*B.subtilis 168*	*E.coli 25922*
GT10-25		>192	>192	>192
GT10-26		192	48	192
GT10-27		48	24	96
GT10-28		>192	>192	>192
GT10-29		>96	48	>96

4.3.2 Molecular docking simulation

The two compounds, **GT10-22** with a *meta*-fluoro and **GT10-27** with a *meta*-nitro substituent on the phenyl ring on R_1 were found to exhibit higher antibacterial activity than the hit compound **GT10** with a *meta*-methyl group. The binding modes of all three compounds are similar (Figure 4.1C). The ICM binding scores for **GT10-22** and **GT10-27** are –33.8 and –34.6 respectively. The guanidinium moiety at the R_2 group of **GT10-22** forms three hydrogen bonds with Tyr155 (3.16 Å) and Ser160 (2.97 Å and 3.02 Å) (Figure 4.1A). The guanidinium moiety at the R_2 group of compound **GT10-27** also forms three hydrogen bonds with Tyr155 (3.16 Å) and Ser160 (2.98 and 3.04 Å) (Figure 4.1B). In addition, the phenyl ring of the isatin core and the 3-nitrophenyl ring of **GT10-27** establish hydrophobic interactions with the side chains of Ile195, Val233 and Pro234 respectively. Apart from forming strong hydrogen bonds and hydrophobic interactions, the molecular geometry of **GT10-22** or **GT10-27** favors its fit into the GT binding pocket of *S. aureus* PBP 2. Compounds **GT10-22** and **GT10-27** were observed to perfectly occupy the upper region of GT binding site where the growing glycan chain should be located based on the proposed mechanism for lipid II polymerization [33]. Furthermore, the *meta*-fluoro and *meta*-nitro substituent groups on R_1 seem to facilitate cell permeability probably due to an increase in polarity of the compounds.

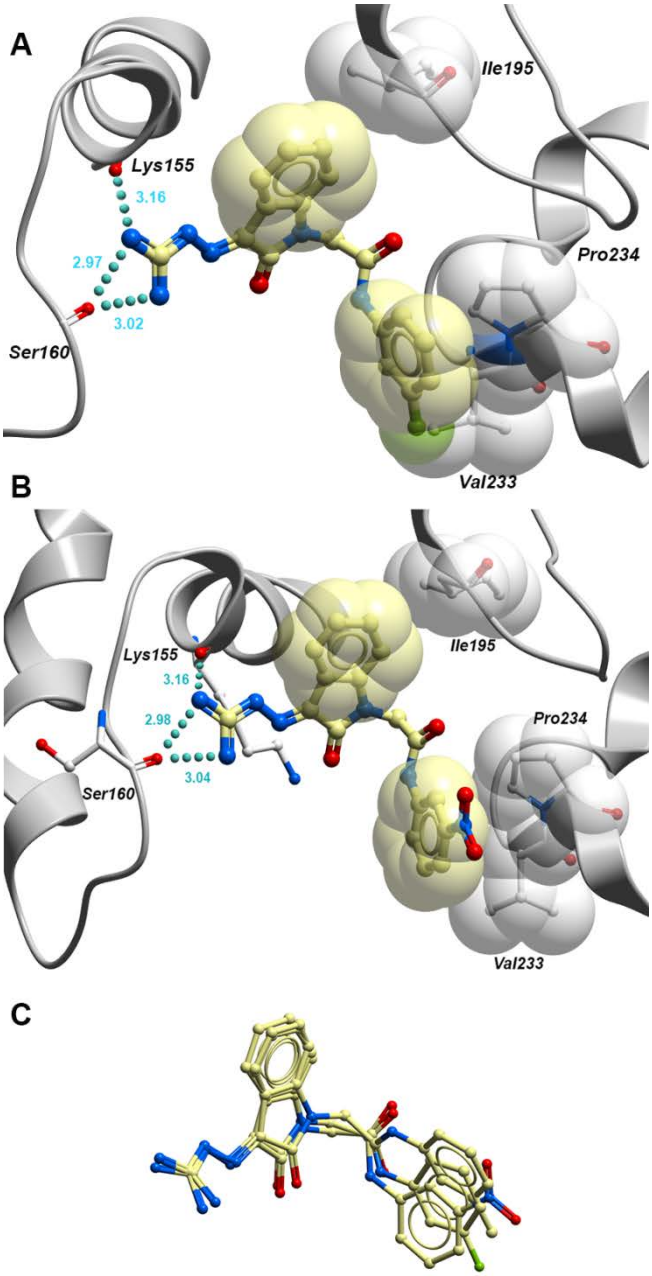


Figure 4.1 (A) Predicted binding pose of **GT10-22** to the GT binding site of *S. aureus* PBP 2. (B) The predicted binding pose of **GT10-27** to the GT binding site of *S. aureus* PBP 2. Hydrophobic residues and hydrophobic moieties of compounds **GT10-22** and **GT10-27** are shown in space-fillling form. Hydrogen bonds are represented with cyan spheres. The labeled interacting residues are shown with a "ball and stick" model and colored by atom types. (C) Superimposition of the predicted binding poses of **GT10**, **GT10-22** and **GT10-27** in the GT binding site.

4.3.3 Saturation-transfer difference (STD)-NMR

It was not difficult to obtain the STD-NMR between **GT10-27** and GT domain due to its strong binding to GT domain and good solubility in buffer solution. The protons on this compound could be precisely assigned in the COSY 2D NMR spectrum (Figure 4.2).

The STD-NMR and competitive STD-NMR spectra between **GT10-27** and GT domain are shown in Figure 4.3. Saturation of GT domain was performed by selective irradiation at -0.4 ppm where no ligand signals were present. Saturation was transferred to protons in close contact with the protein by intermolecular saturation transfer. The STD with saturation spectrum (Figure 4.3c) confirmed that **GT10-27** binds to the GT domain, since saturation on all aromatic protons was observed. The proton signals on methylene overlapped with the water signal (not shown in spectra). The spectrum also shows that proton H_1 has the highest degree of saturation (set to 100 %), and other protons (H_2-H_8) have a range of relative degree of saturation from 53 % to 78 %. This observation indicated that there is a binding contact between **GT10-27** and GT domain of *S. aureus* PBP 2. Moreover, the 3-nitrophenyl ring of **GT10-27** is closer to the GT binding site than the isatin core. By contrast, there is no STD-NMR signal in the control STD spectrum without saturation (Figure 4.3b), in which no irradiation power was applied to the protein. To gain insight into the binding mode of **GT10-27** to GT domain, competitive STD-NMR was performed with addition of increasing amounts of high binding-affinity ligand

moenomycin A to a 100:1 complex of **GT10-27**/GT domain. Difference spectra were monitored for a change in intensity of signals belonging to **GT10-27** (H_1-H_8) during the titration. As shown in the Figure 4.3d and 4.3e, addition of 5 μM of moenomycin A results in a uniform decrease in the intensity of signals (H_1-H_8). These results indicated that **GT10-27** and moenomycin A very likely bind to the GT domain at the same location in a competitive manner.

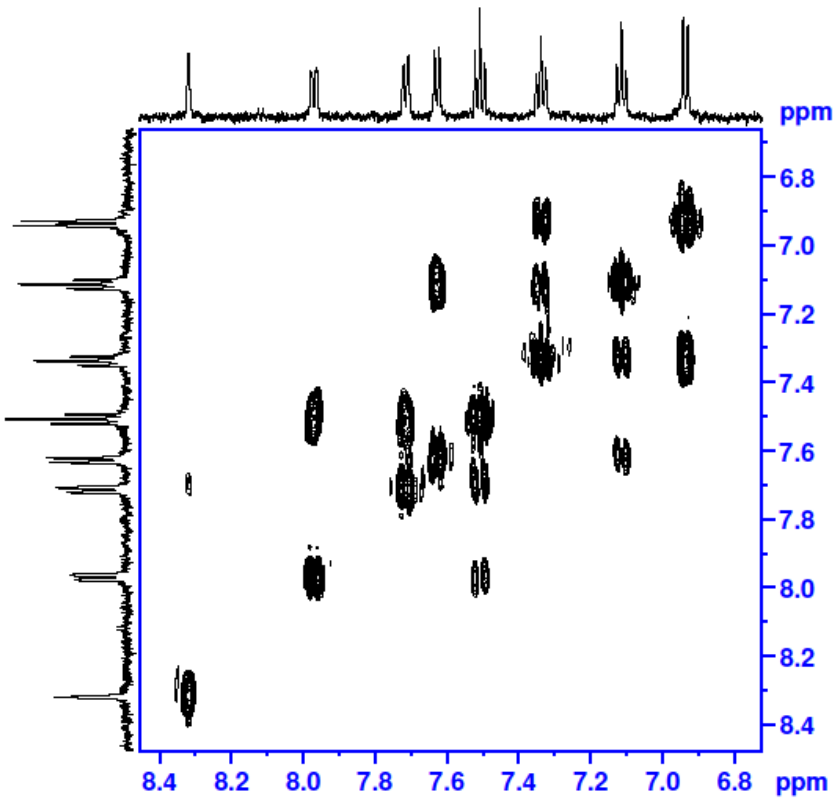


Figure 4.2 COSY 2D NMR spectrum of **GT10-27**.

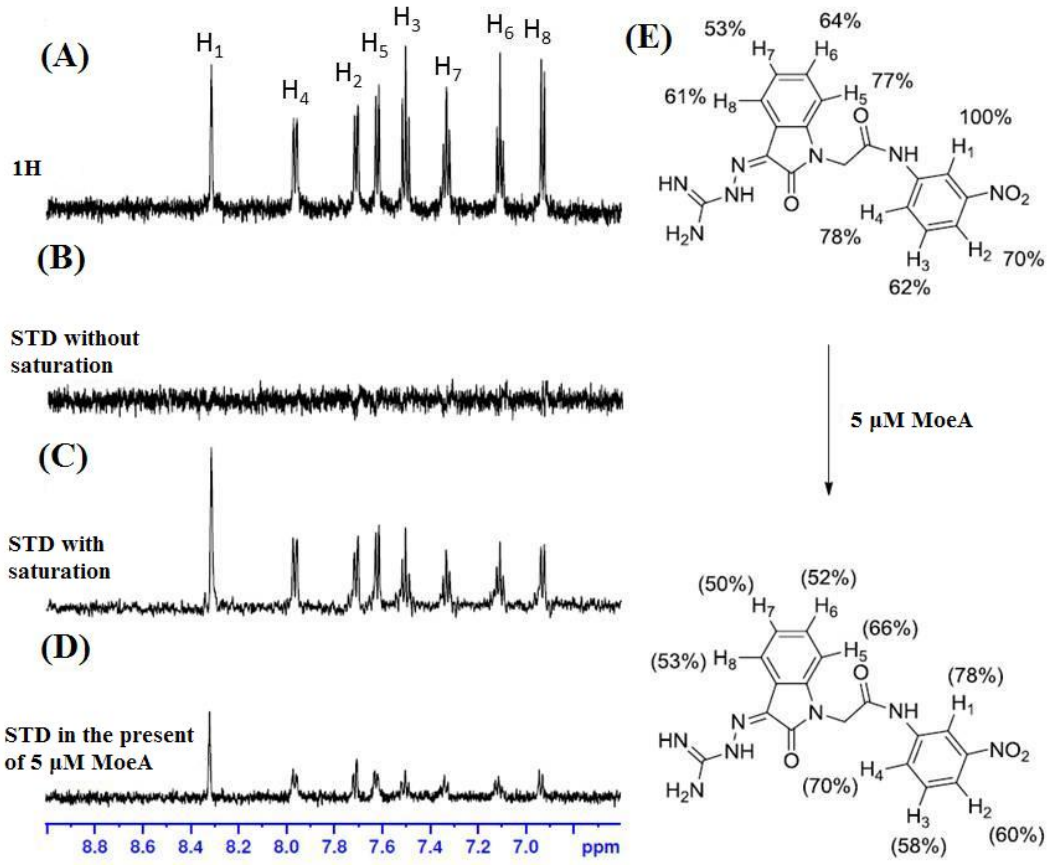


Figure 4.3 STD-NMR study of interaction of **GT10-27** with GT domain of *S. aureus* PBP 2. **a)** ^{1}H-NMR of **GT10-27**. **b)** STD-NMR spectrum without saturation **c)** STD-NMR spectrum between **GT10-27** and GT domain. 2.5 µM GT domain, 250 µM **GT10-27** in 20 mM sodium phosphate buffer in D_2O, 5 % DMSO-d_6. **d)** STD-NMR spectrum in the presence of 5 µM moenomycin A. **e)** Relative degrees of saturation of the individual protons on **GT10-27** with and without 5 µM moenomycin A, H_1 was set to 100 %.

4.3.4 Cytotoxicity

The cytotoxicity effect of **GT10-27** against three different mammalian cell lines was measured. According to the results summarized in Table 4.2, this compound exerts toxic effect on the mammalian cells, especially the normal skin cells, since it inhibits the proliferation of Detroit 551 and HFF1 at low concentrations (IC_{50} < 10 μM). The cytotoxicity effect may be attributed to the nitro-aromatic structure which is difficult to be oxidatively degraded, as the nitro group is stabilized by the phenyl ring [132]. Furthermore, the isatin core structure has also been reported to be cytotoxic [133].

Selectivity index (SI), the ratio between IC_{50} and MIC, is commonly used to evaluate whether a compound can be a good drug candidate. Ten is always used as the threshold [134]. However, the SI value of **GT10-27** against *B.subtilis* is only 0.13, which suggests that this compound is strongly cytotoxic. Thus, further modification should be applied to this compound before it can be considered as a drug lead.

Table 4.2 Cytotoxicity of GT10-27

Compound	Cytotoxicity (IC50, μM)		
	A375	Detroit 551	HFF1
GT10-27	25.00	7.99	6.96

A375: human melanoma cell line;

Detroit 551: normal human skin fibroblast;

HFF1: normal human foreskin fibroblast.

4.4 Concluding remarks

In this chapter, the structure activity relationship of **GT10** analogues was established. The results indicated that when the phenyl ring in R_1 is substituted by functional groups such as nitro, fluorine or ether at the *meta*-position, the antibacterial ability is improved. The MIC of the most active compound **GT10-27** is 24 μg mL^{-1} against *B. subtilis* which is 8 times higher than parent compound **GT10**. Besides, the guanidyl group in R_2 is also essential to the activity.

Molecular modeling of compounds **GT10-27** to the known GT structure shows a binding interface which partially overlaps with that of moenomycin A. The docking results described in this chapter, together with results from STD-NMR and antibacterial susceptibility assays, provide useful information that will enable further chemical optimization of **GT10-27** into potent GT inhibitors against antibiotic-resistant bacteria.

Chapter 5

Conclusions

The wide spread of drug-resistant bacteria has been a huge threat to human health. Thus, there is an urgent need to discover novel antibacterial agents with new drug targets. Glycosyltransferase (GT), which catalyzes the formation of bacterial cell wall, is a promising target for the development of new antibiotics. The structure of GT is highly conserved in both wild-type and drug resistant strains and it has no counterpart in the eukaryotic cells. In this thesis, molecular docking, a powerful tool for drug discovery, was applied to screen out the hit compounds that can potentially be used as glycosyltransferase inhibitor. After screening drug databases containing over 3,000,000 compounds against the high resolution X-ray structure of the GT domain of *S. aureus* PBP 2 by ICM, compound **GT10**, one of 34 top-ranking structures, showed a weak antibacterial activity against *S. aureus* and *B. subtilis* (192 μg mL^{-1}). Also, the interaction between this compound and the GT domain was confirmed by STD-NMR.

In order to optimize this hit compound, a series of its analogues were designed and synthesized, and a structure activity relationship was proposed according to their MICs results against *E. coli*, *S. aureus* and *B. subitlis*. The results indicated that the phenyl ring on R_1 substituted by functional groups such as nitro, fluorine or ether groups at the *meta*-position improved the antibacterial ability. Substitution on the other sites or disubstitution would not enhance the antibacterial ability. When this phenyl group was replaced by biphenyl or substituted naphthyl, the antibacterial ability was reduced

probably because they are too bulky to enter the binding site. Besides, the guanidyl group in R_2 is also essential to the antibacterial activity since it may bind firmly with the GT domain through hydrogen bonds. **GT10-27** was found to be the most active compound and its MIC values against *E. coli, S. aureus* and *B. subitlis* were 96, 48, 24 μg mL^{-1},respectively.

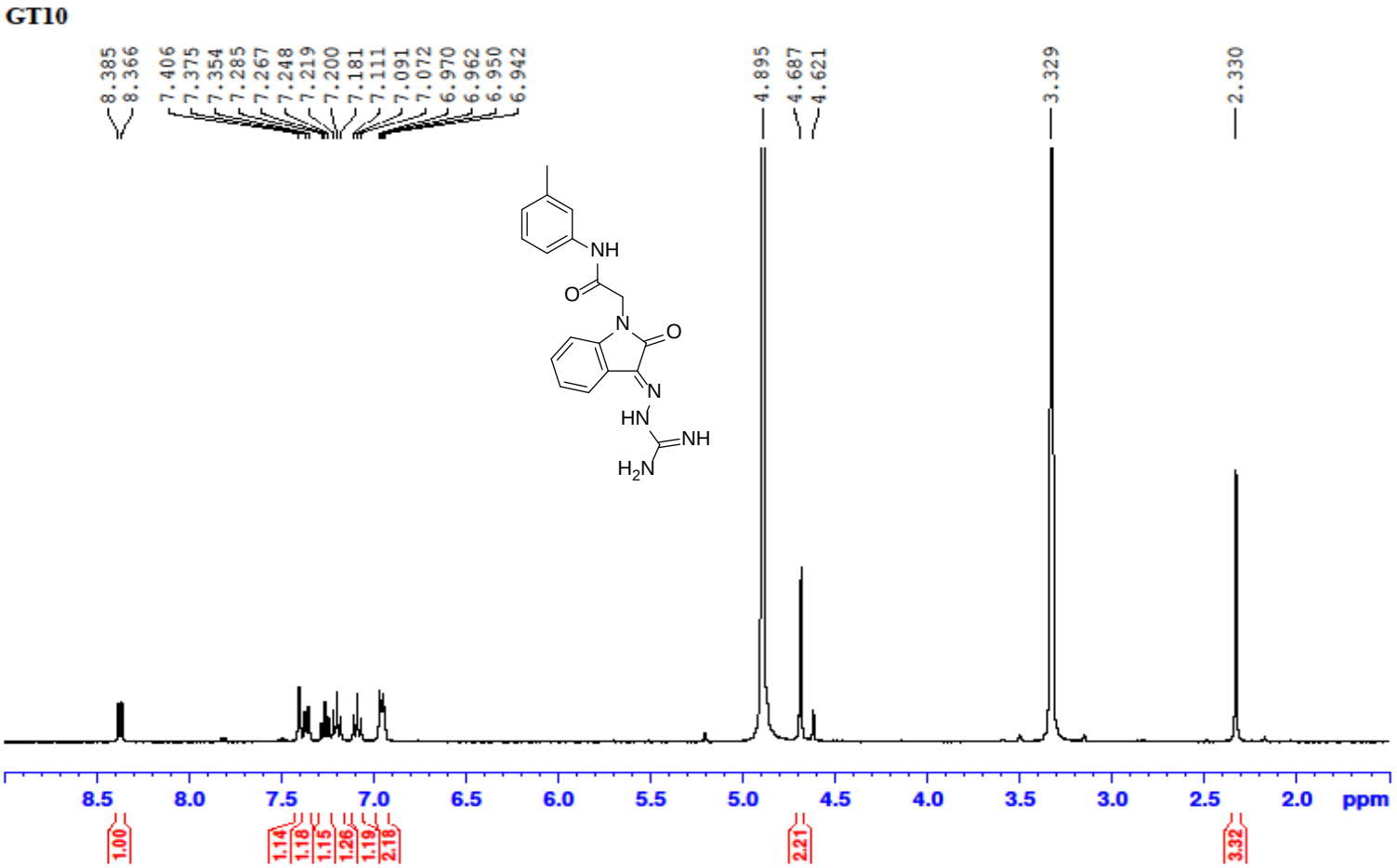
Molecular docking was used to simulate the binding pose between the active **GT10-27** and the GT domain of PBP 2. The results of modeling revealed that this compound binds firmly with the GT domain with an ICM binding score of -34.6. It was also observed that the guanidyl group in the compound forms three hydrogen bonds with Tyr155 and Ser160. Moreover, the phenyl moieties in the isatin core or in R_1 contribute to the strong binding by hydrophobic interaction with Ile195, Val233 and Pro234. STD-NMR results confirmed the interaction between **GT10-27** and GT domain. Competitive STD-NMR results indicated that **GT10-27** and moenomycin A share the same binding site in the GT domain. All these results suggested that **GT10-27** could be a potential glycosyltransferase inhibitor.

The results presented in this thesis are very encouraging. However, further work on the hit optimization under the guide of molecular docking is still needed. More **GT10** derivatives should be designed and synthesized to improve the antibacterial ability and reduce the cytotoxic effect to mammalian cells. The hydrophobic moiety in R_1 could be enlarged (but not larger than a naphthyl ring) to enhance the hydrophobic interaction with GT

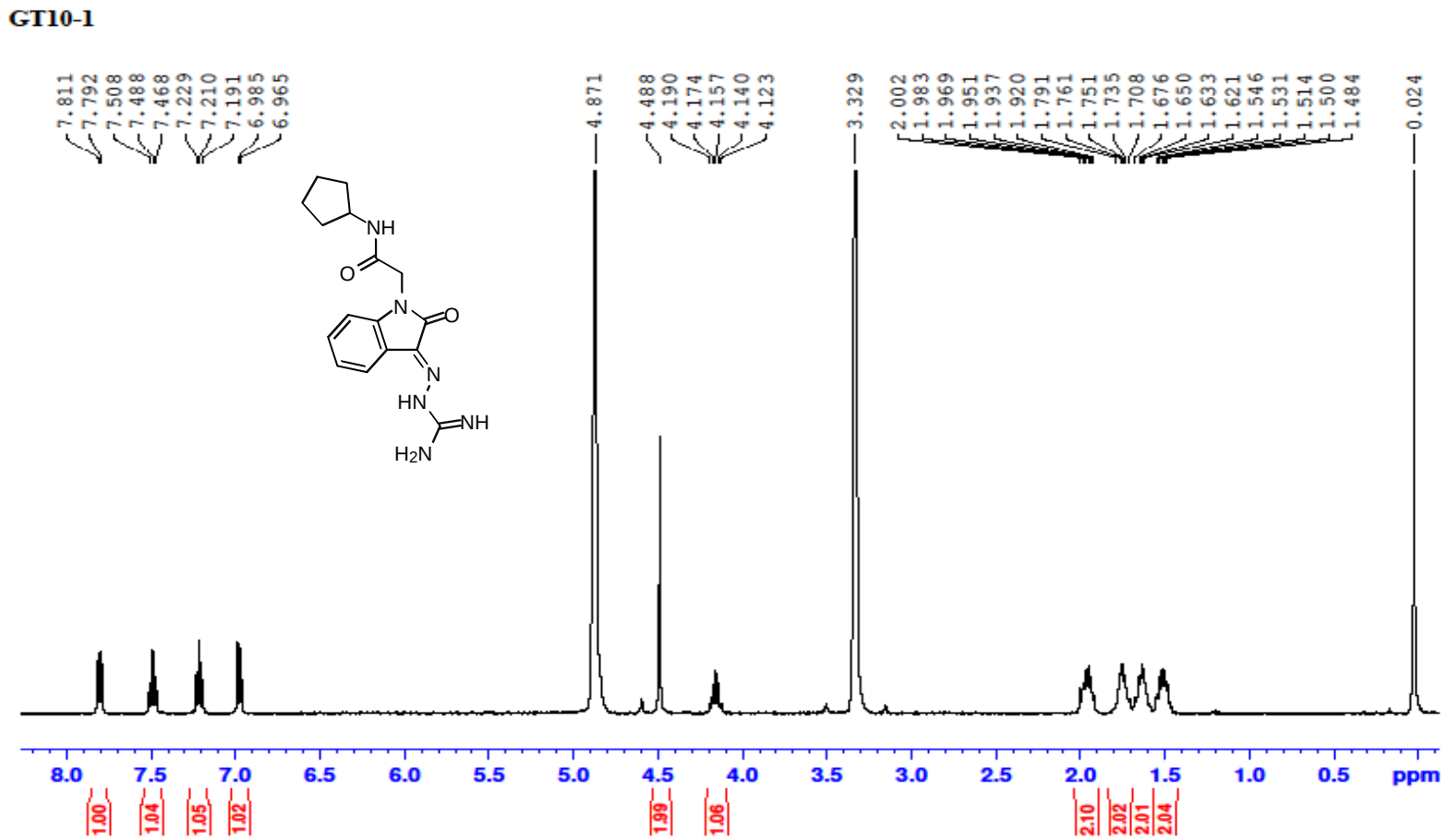
domain. Substitution with different functional groups such as benzheterocycles could be performed to gain stronger binding affinity. Besides, the design of new **GT10** derivatives should follow the Lipinski's rule of five to avoid the membrane permeability problem [135]. It may be feasible to make substitution with electron donating groups such as methyl group, methoxyl group or hydroxyl group on different sites of the isatin core too, since this moiety shows hydrophobic interaction with the GT domain. Besides, the substituted functional groups may reduce the cytotoxicity by favoring the metabolic process of these compounds in cell. To investigate the precise binding sites between **GT10-27** and the GT domain, mutants of GT domain could be made by site-directed mutagenesis on the basis of their binding pose. Through comparing the STD-NMR results between **GT10-27** with the mutated and wild-type GT domain, amino acid residues that **GT10-27** binds to can be found out. Finally, X-ray crystallography can be applied to investigate the binding between **GT10** derivatives and GT domain in details. The results can also guide the design of new structures as GT inhibitors. Through virtual screening, organic synthesis and bioassays, lead compounds with improved potency and lower cytotoxicity could be identified.

Appendix I

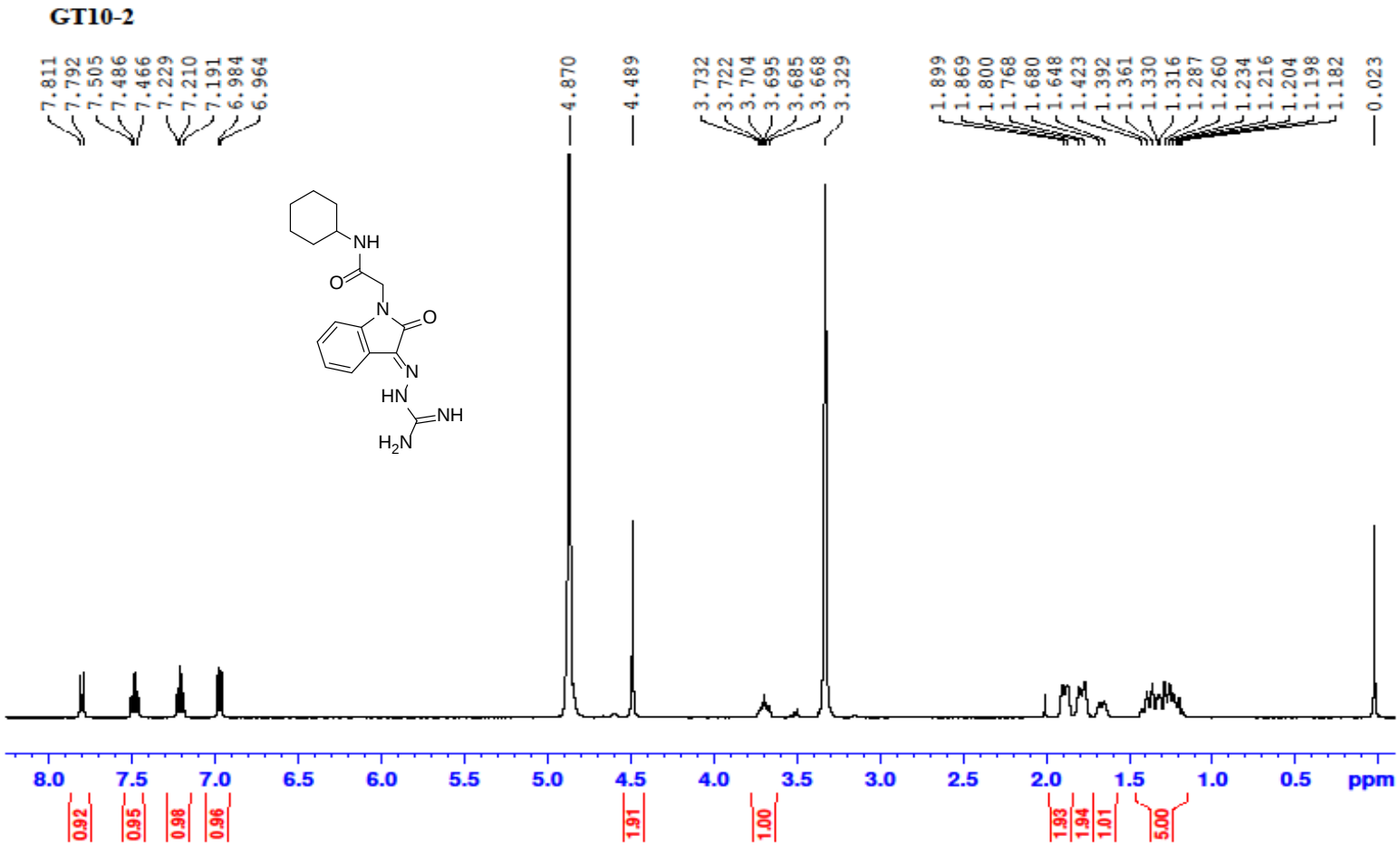
^{1}H-NMR spectra of GT10 and its analogues



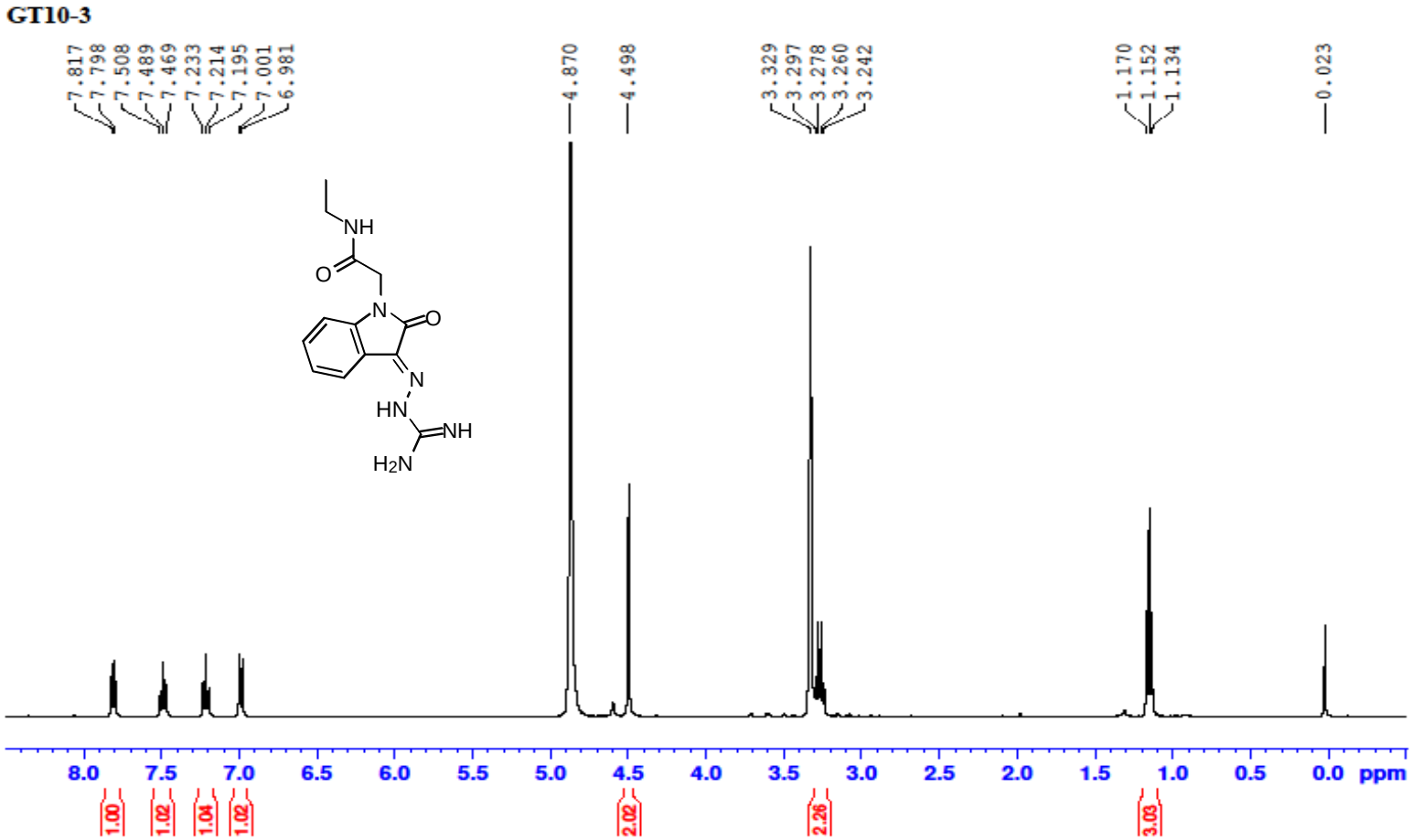
I-1 ^{1}H NMR spectrum of compound GT10



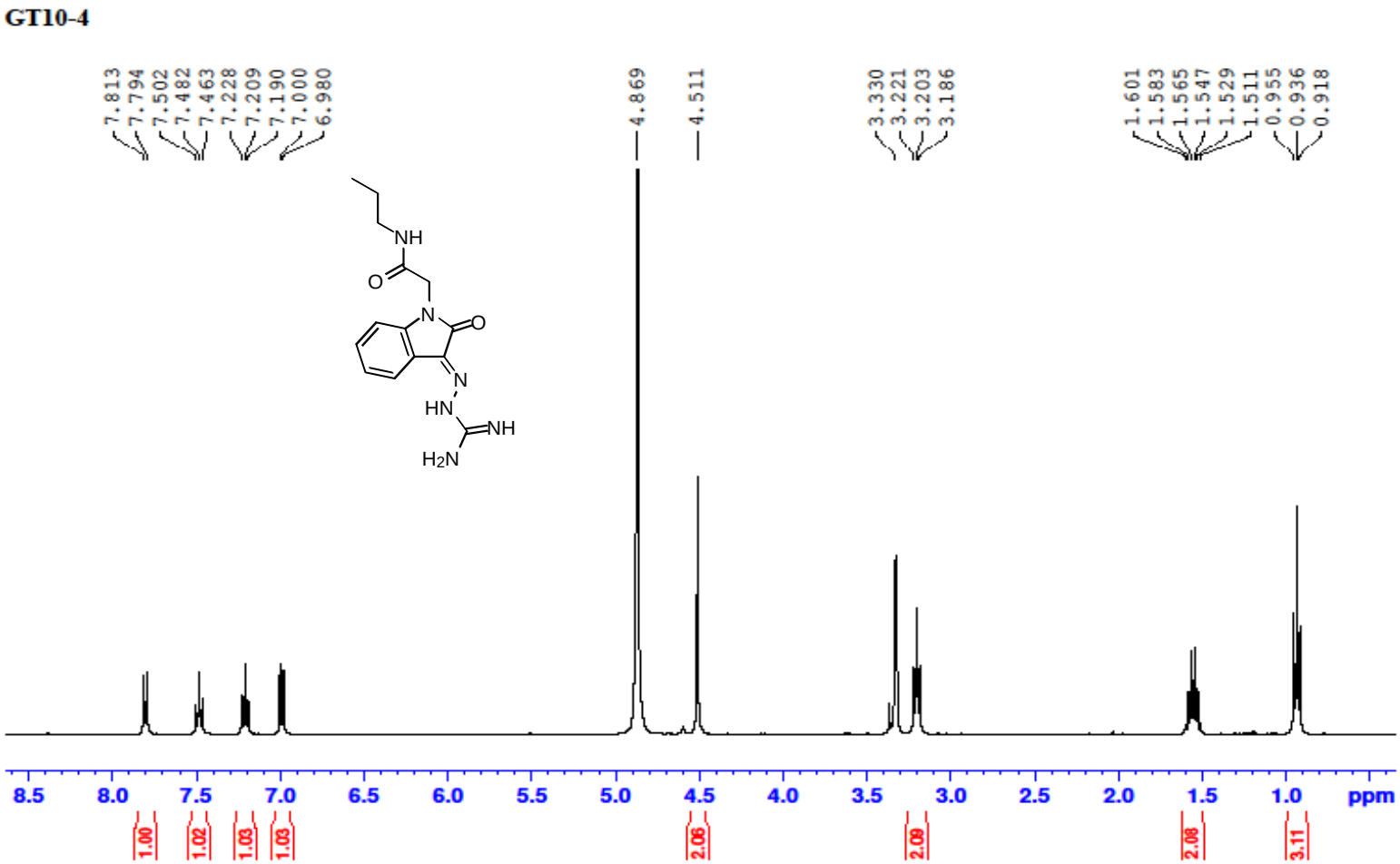
I-2 ^{1}H NMR spectrum of compound GT10-1



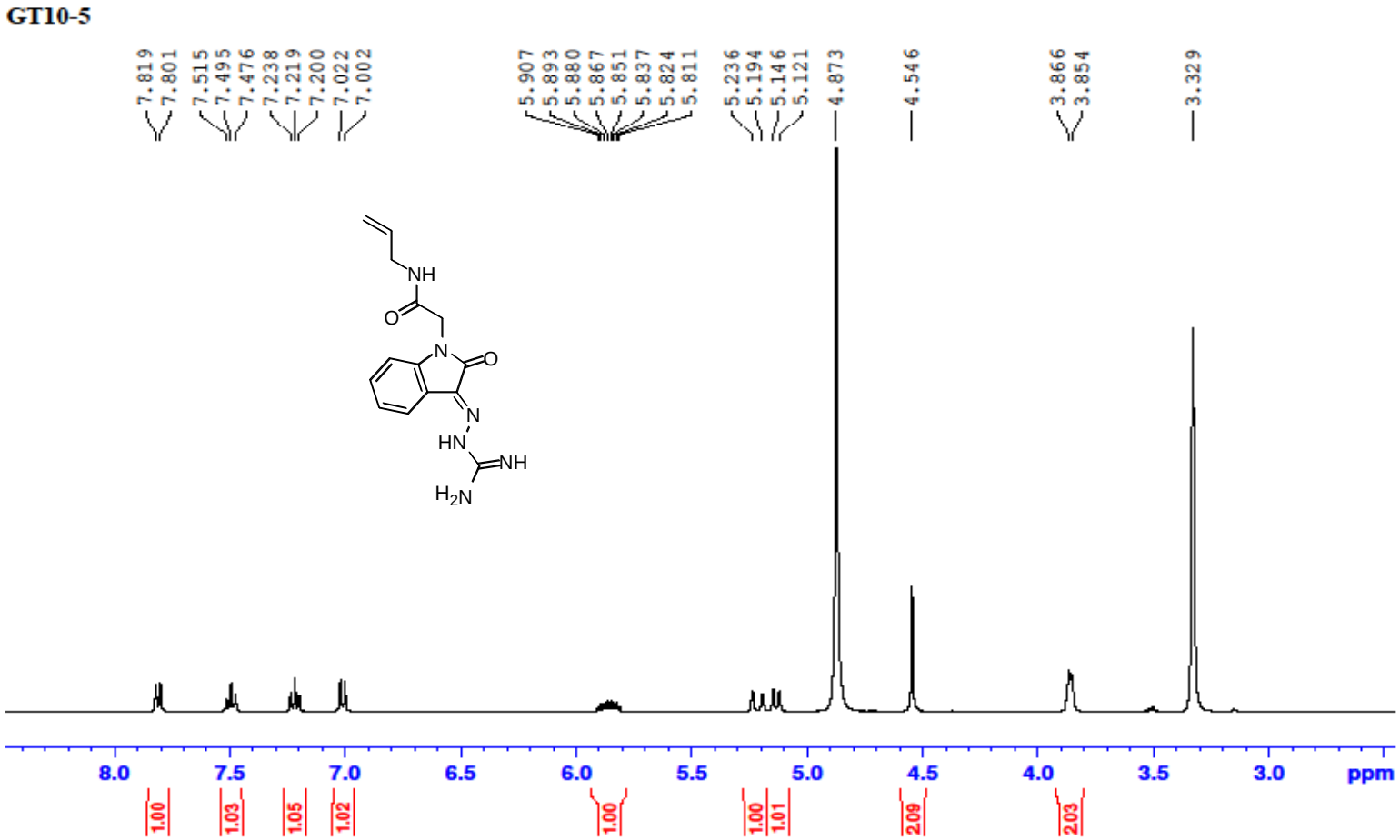
I-3 ^{1}H NMR spectrum of compound GT10-2



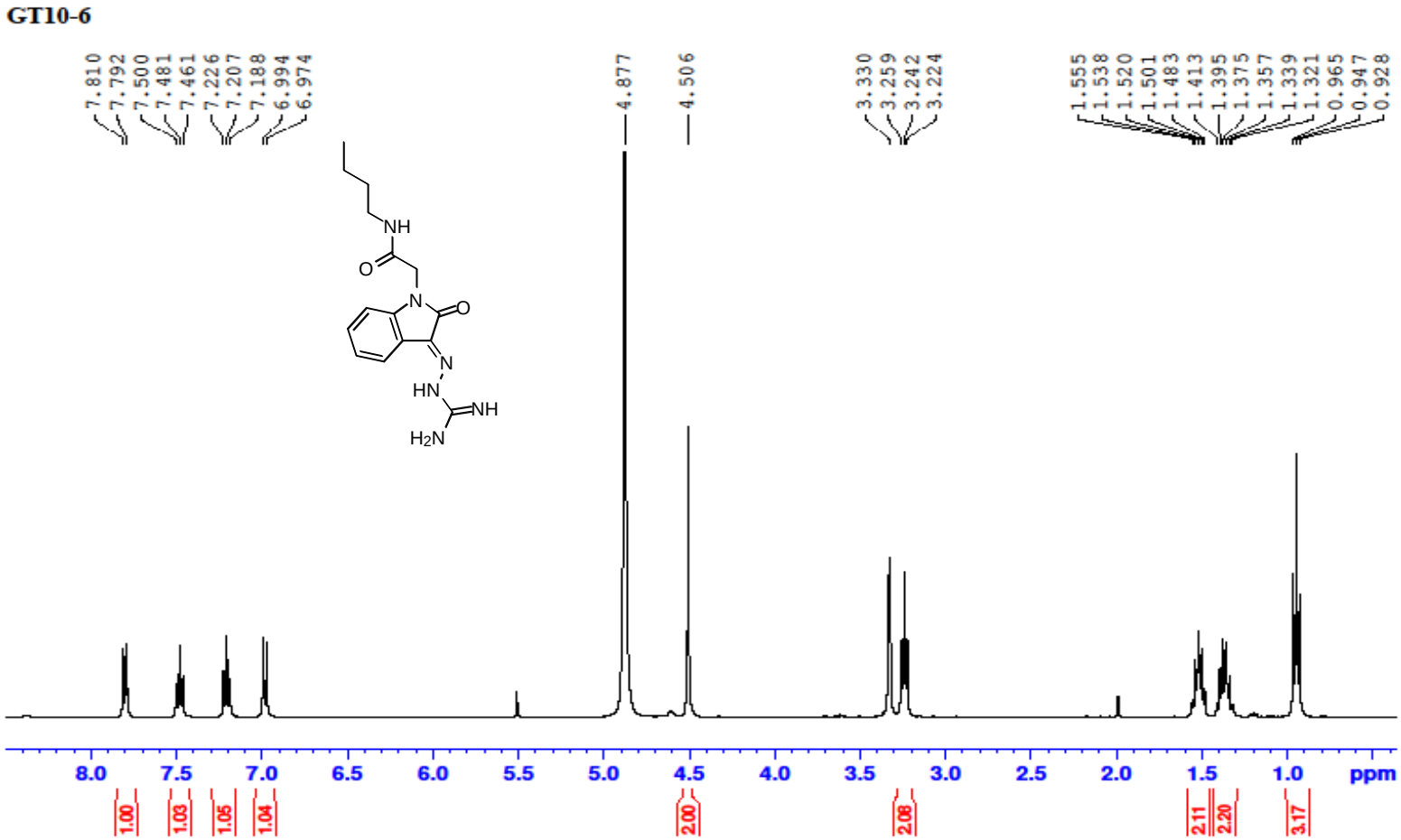
I-4 ^{1}H NMR spectrum of compound GT10-3



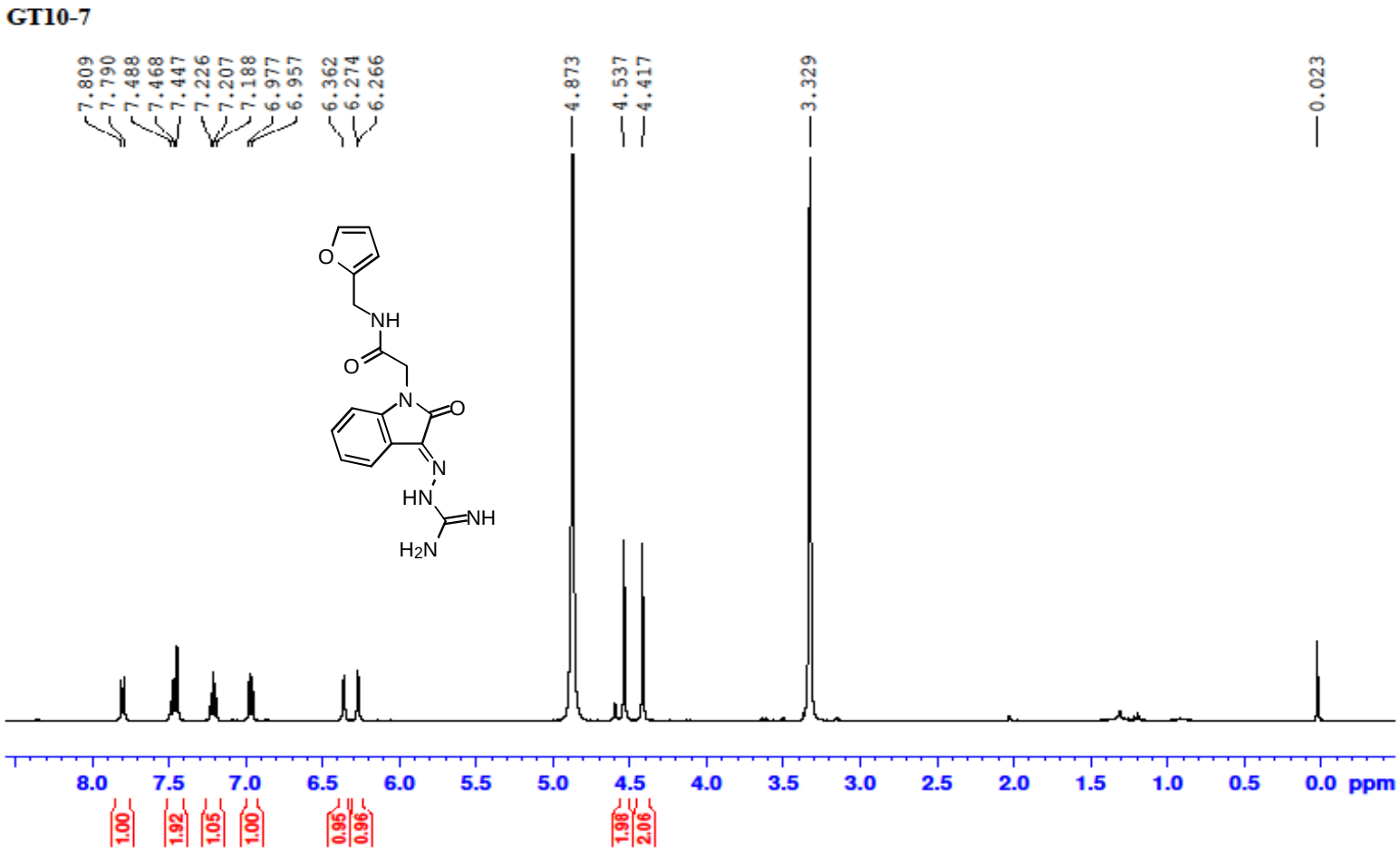
I-5 ^{1}H NMR spectrum of compound GT10-4



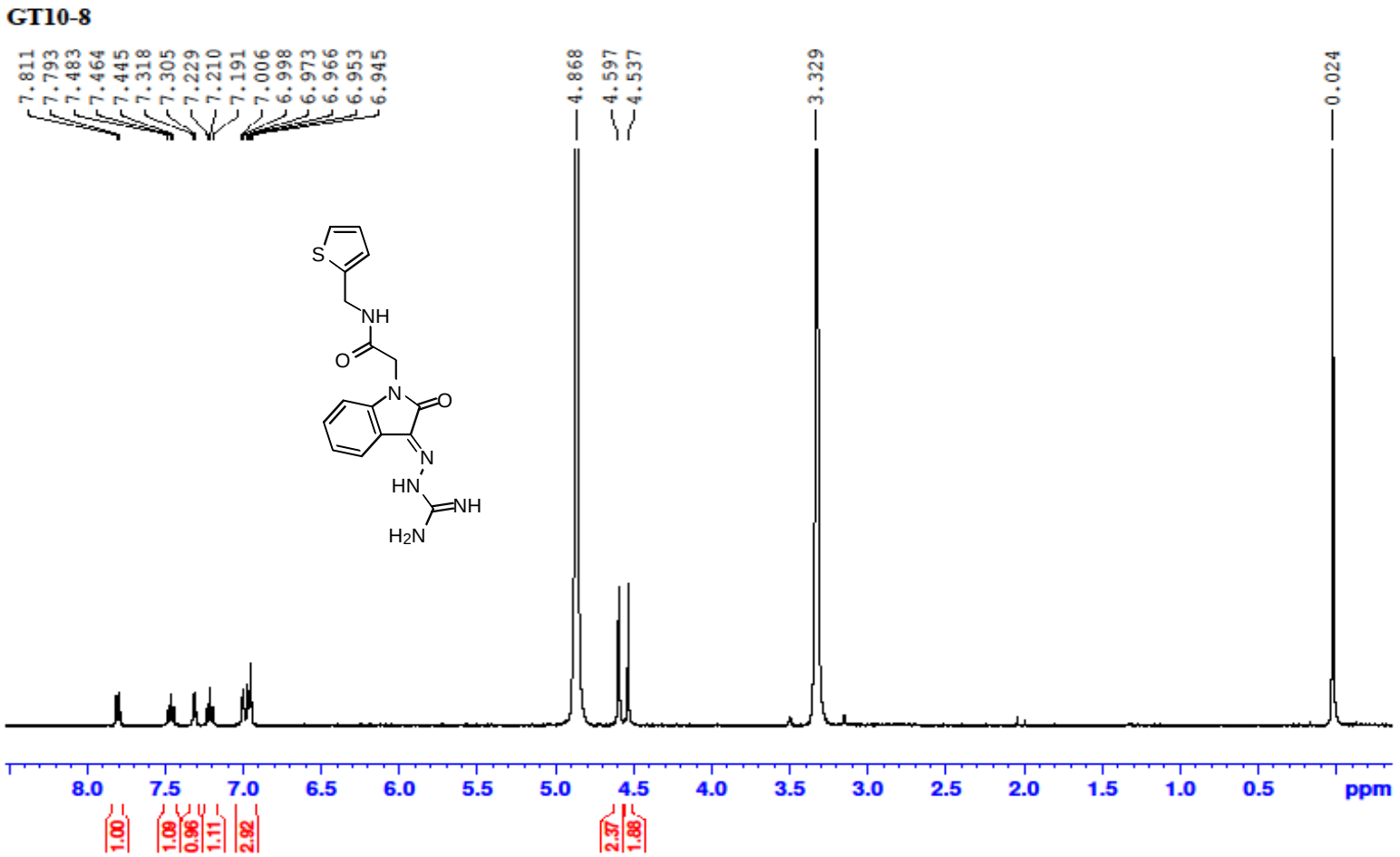
I-6 ^{1}H NMR spectrum of compound GT10-5



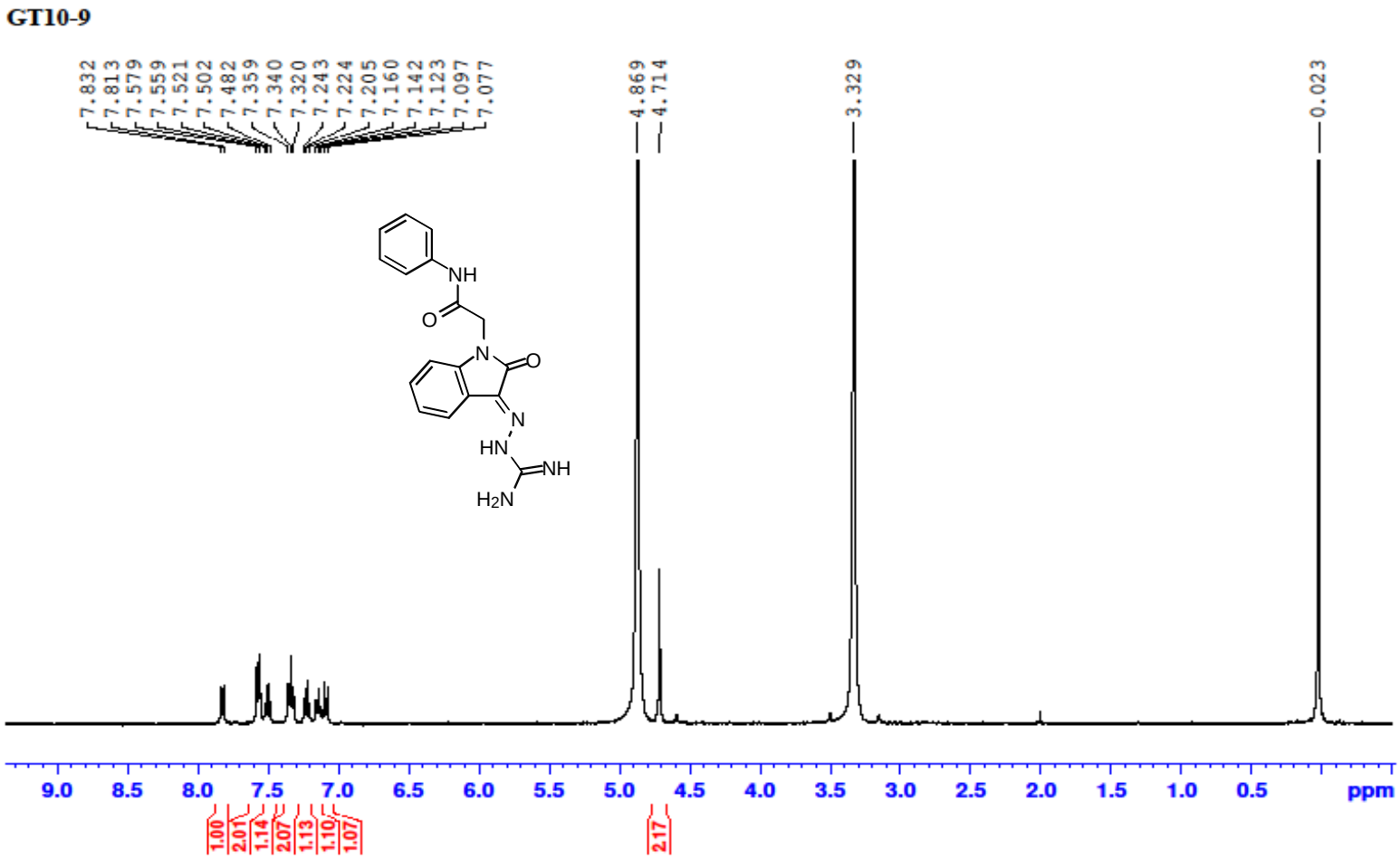
I-7 ^{1}H NMR spectrum of compound GT10-6



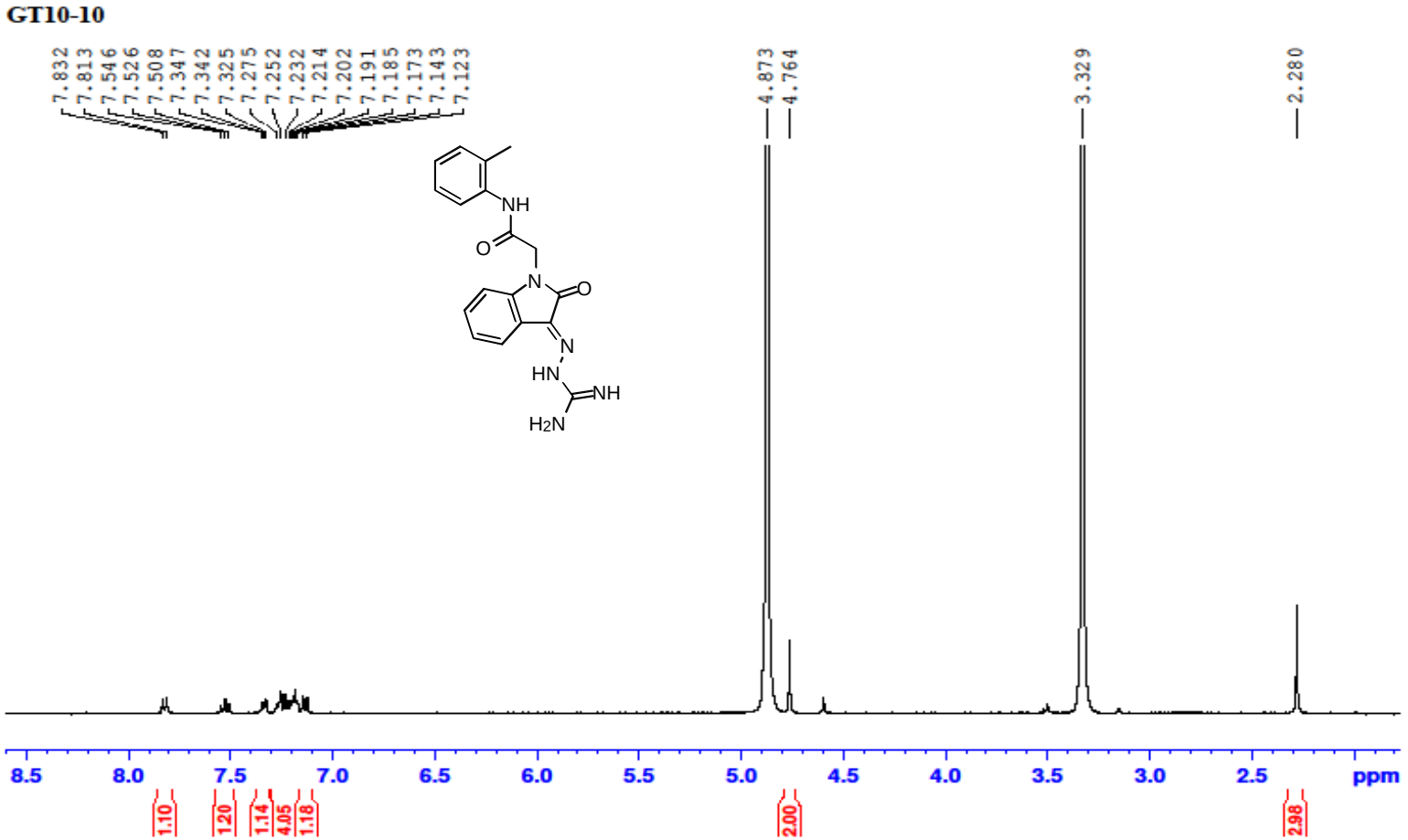
I-8 ^{1}H NMR spectrum of compound GT10-7



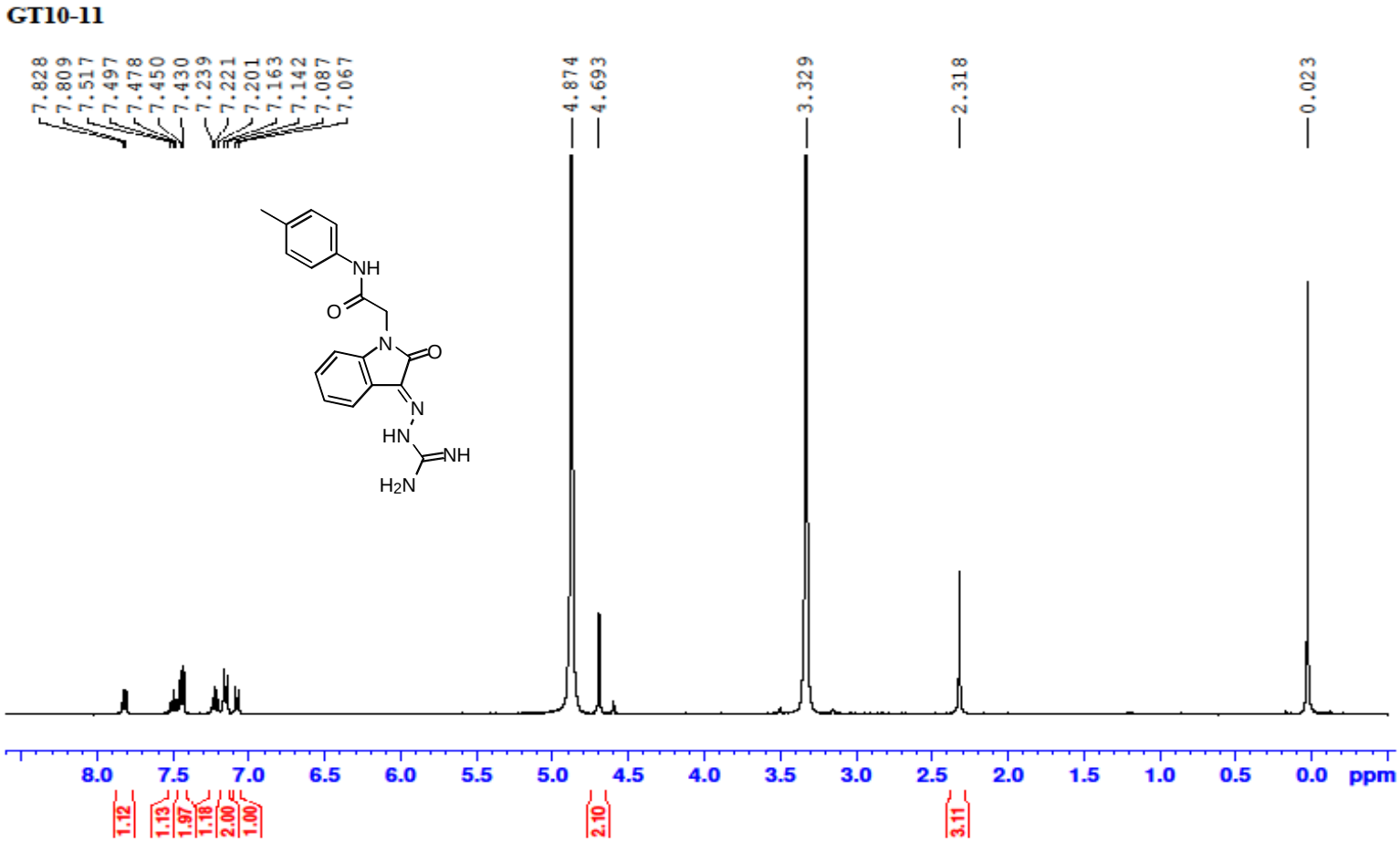
I-9 ^{1}H NMR spectrum of compound GT10-8



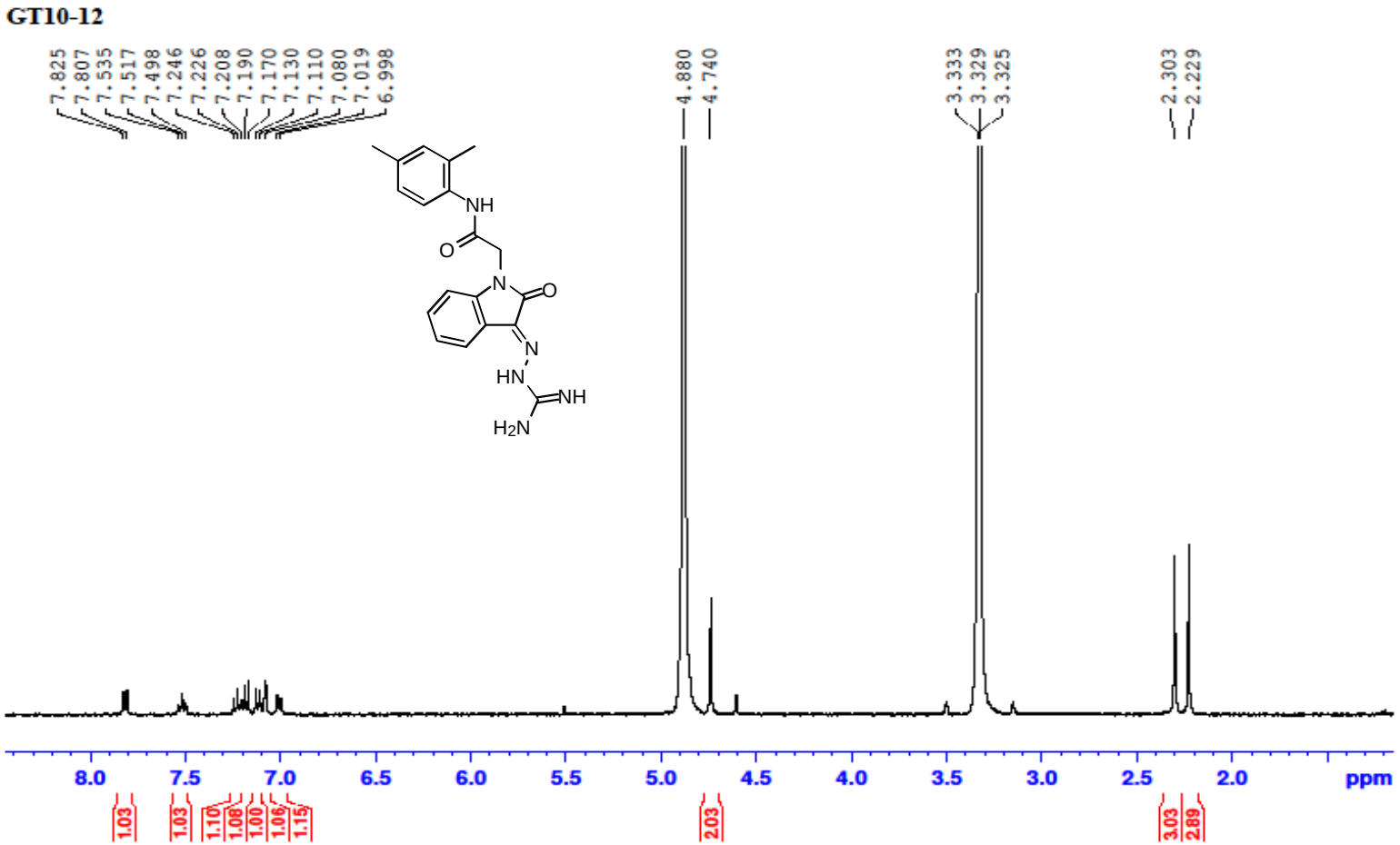
I-10 ^{1}H NMR spectrum of compound GT10-9



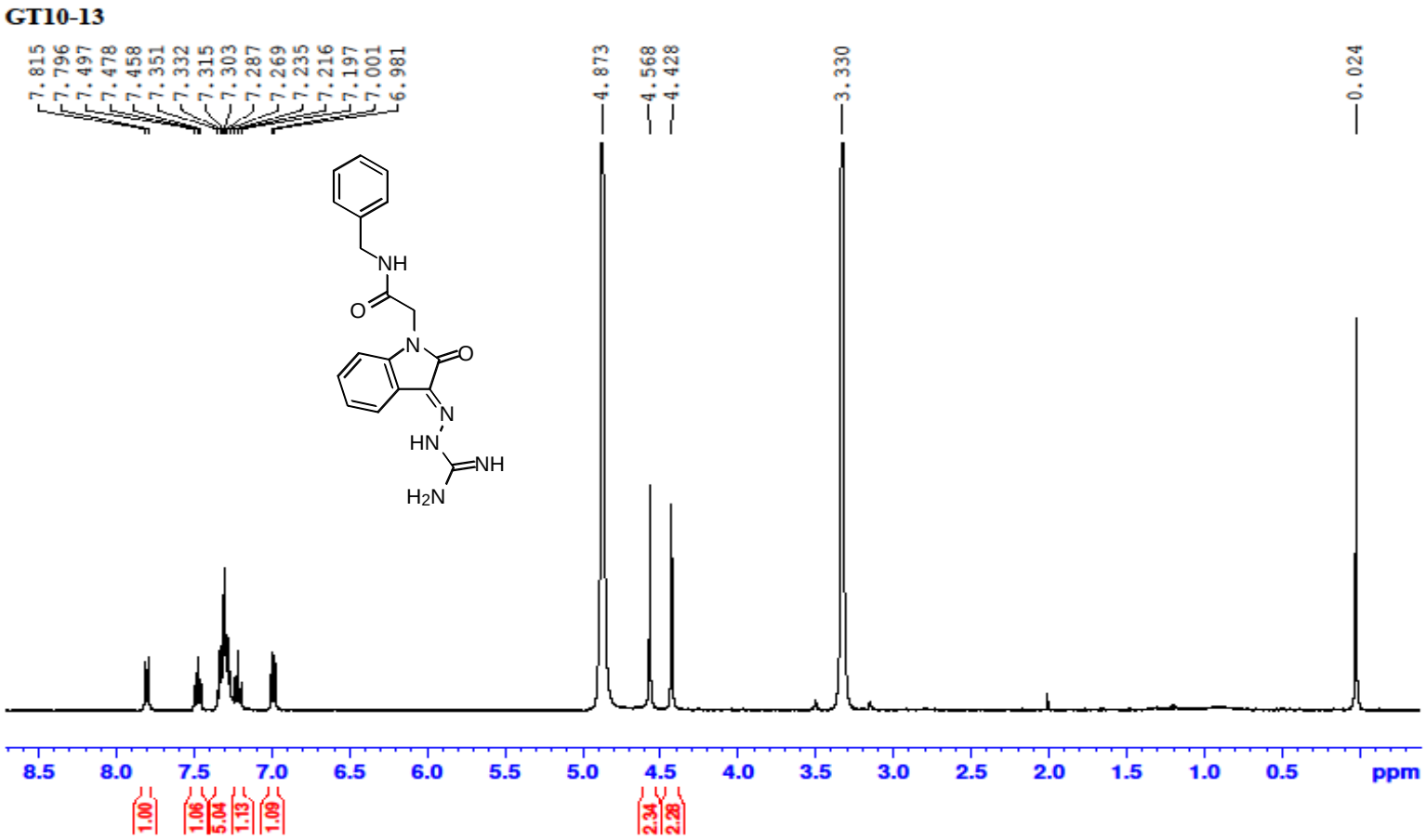
I-11 ^{1}H NMR spectrum of compound GT10-10



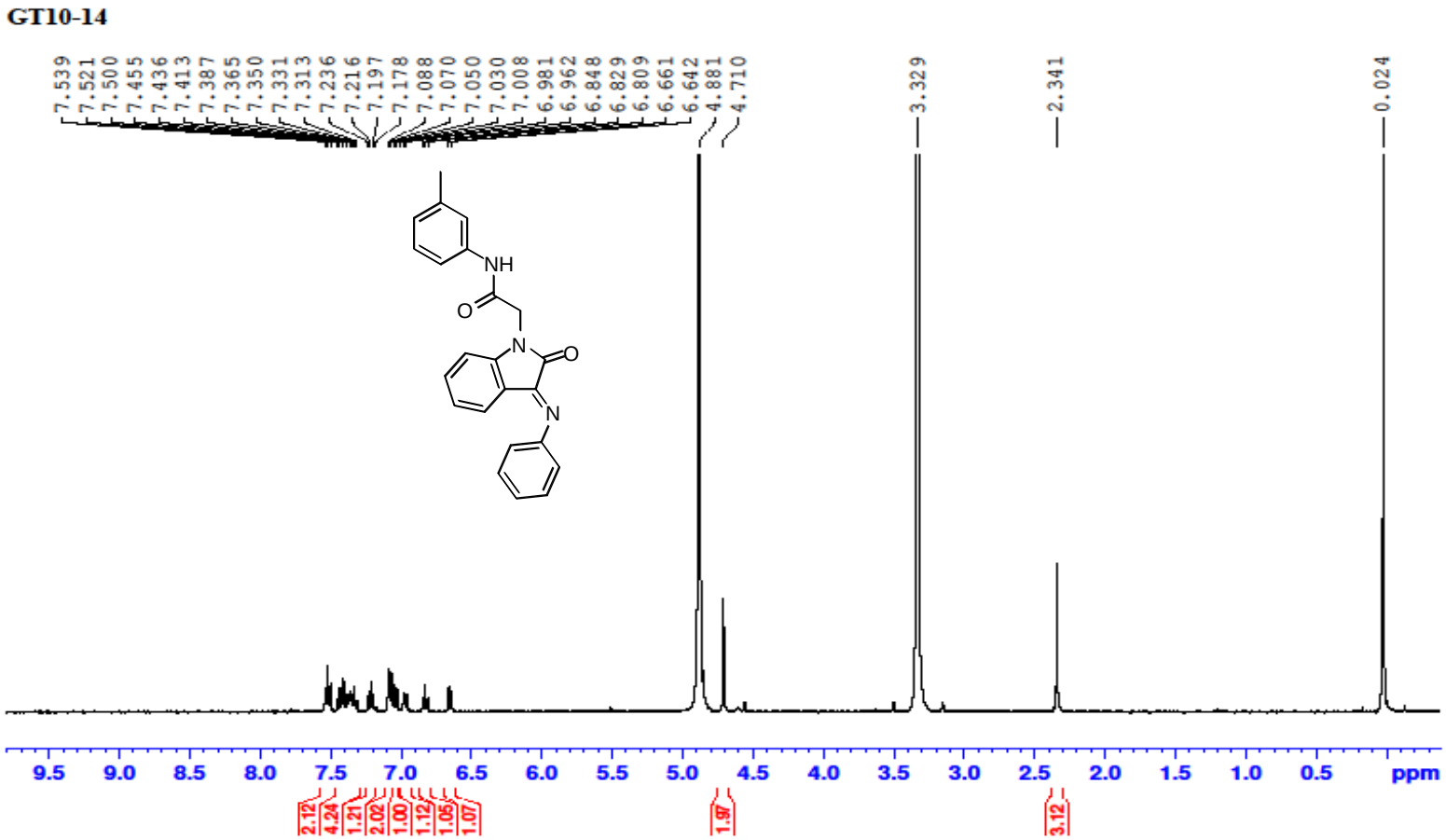
I-12 ^{1}H NMR spectrum of compound GT10-11



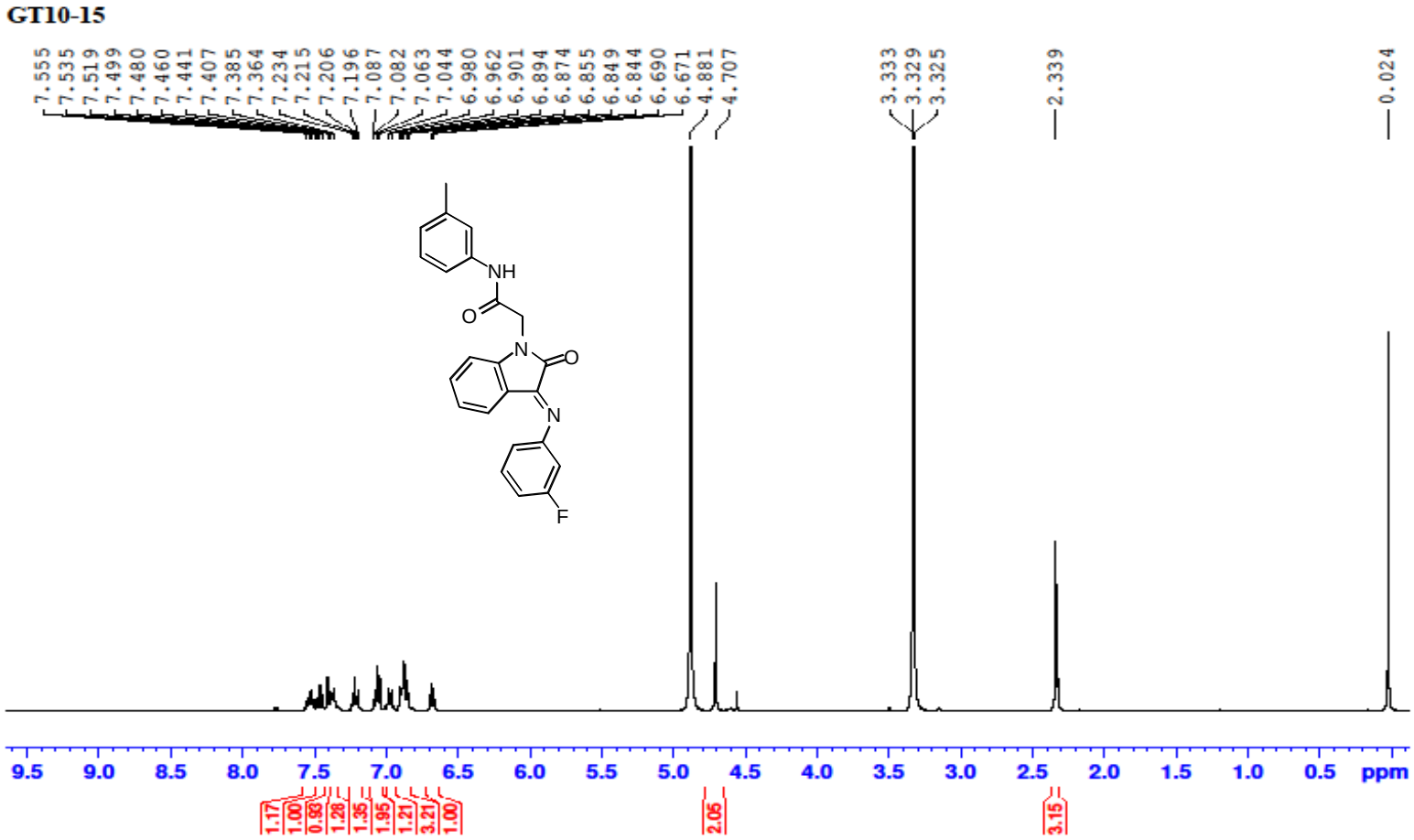
I-13 ^{1}H NMR spectrum of compound GT10-12



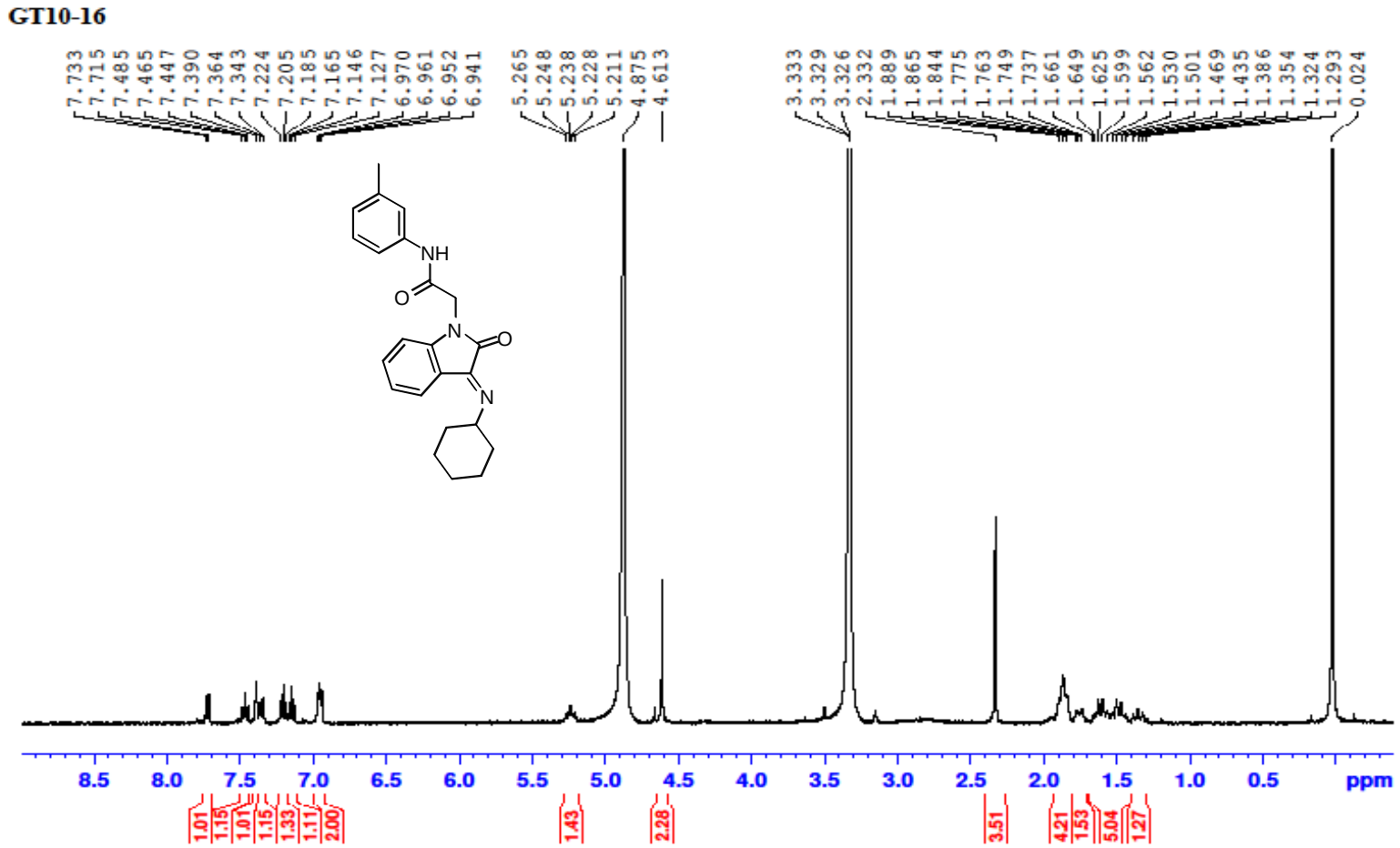
I-14 ^{1}H NMR spectrum of compound GT10-13



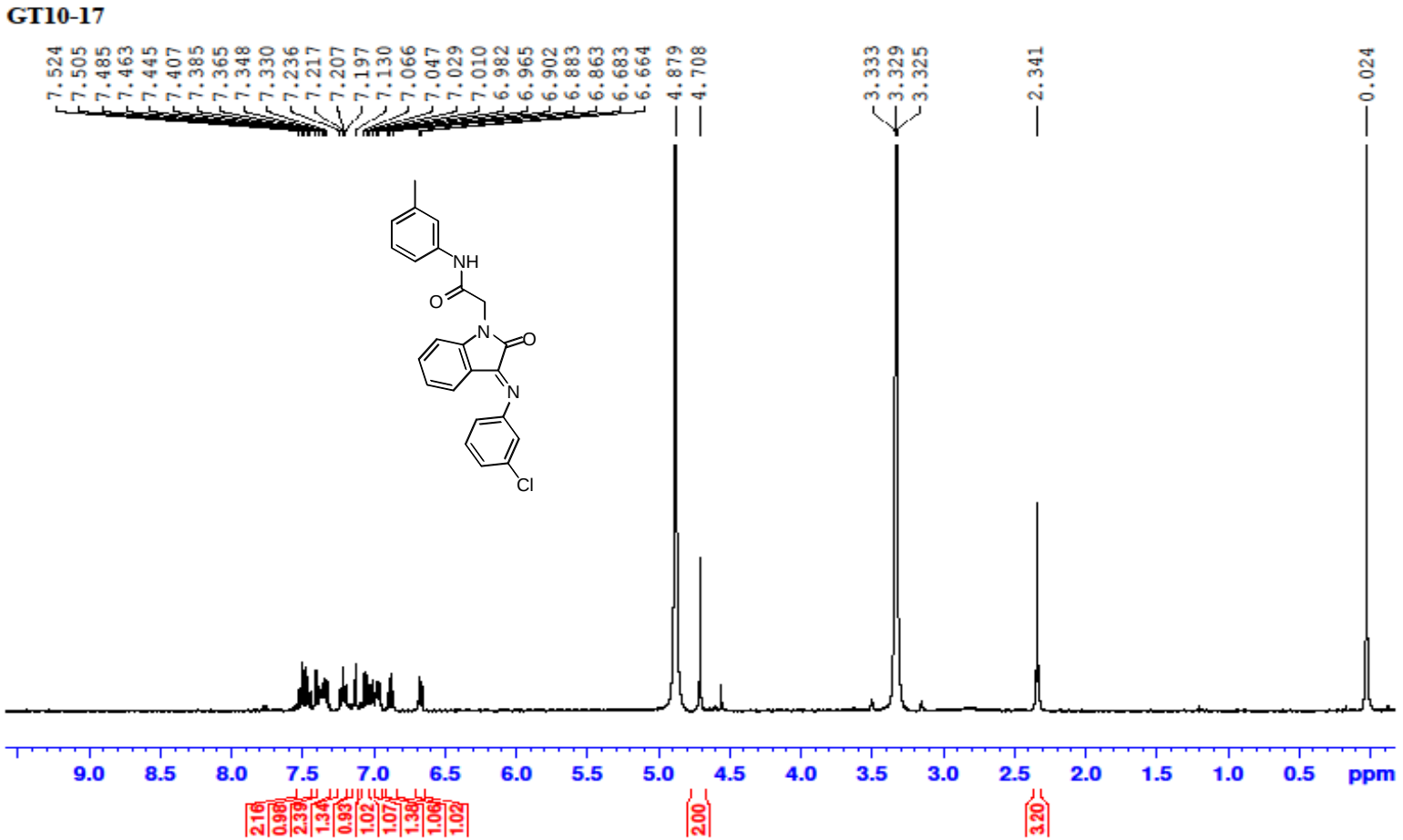
I-15 ^{1}H NMR spectrum of compound GT10-14



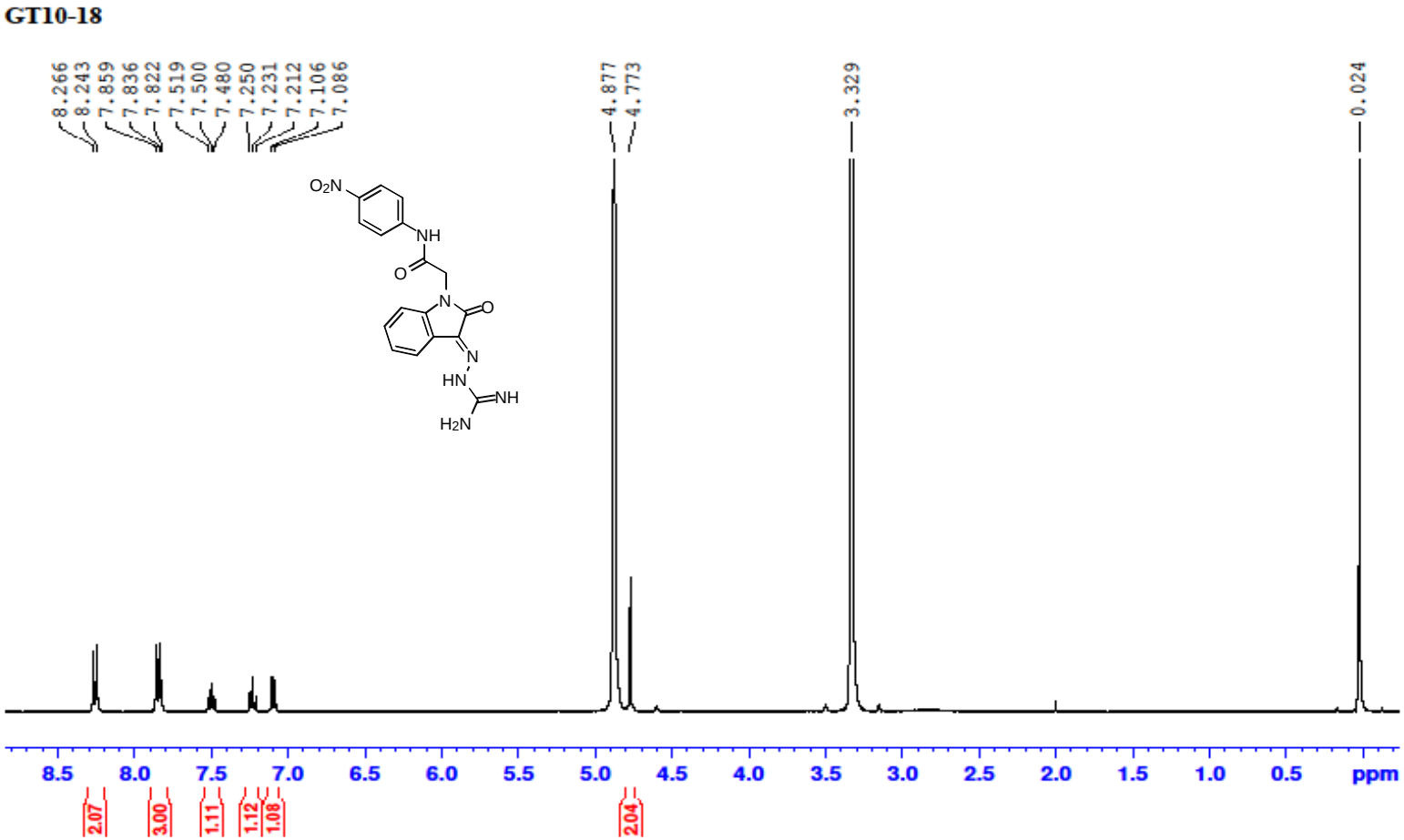
I-16 ^{1}H NMR spectrum of compound GT10-15



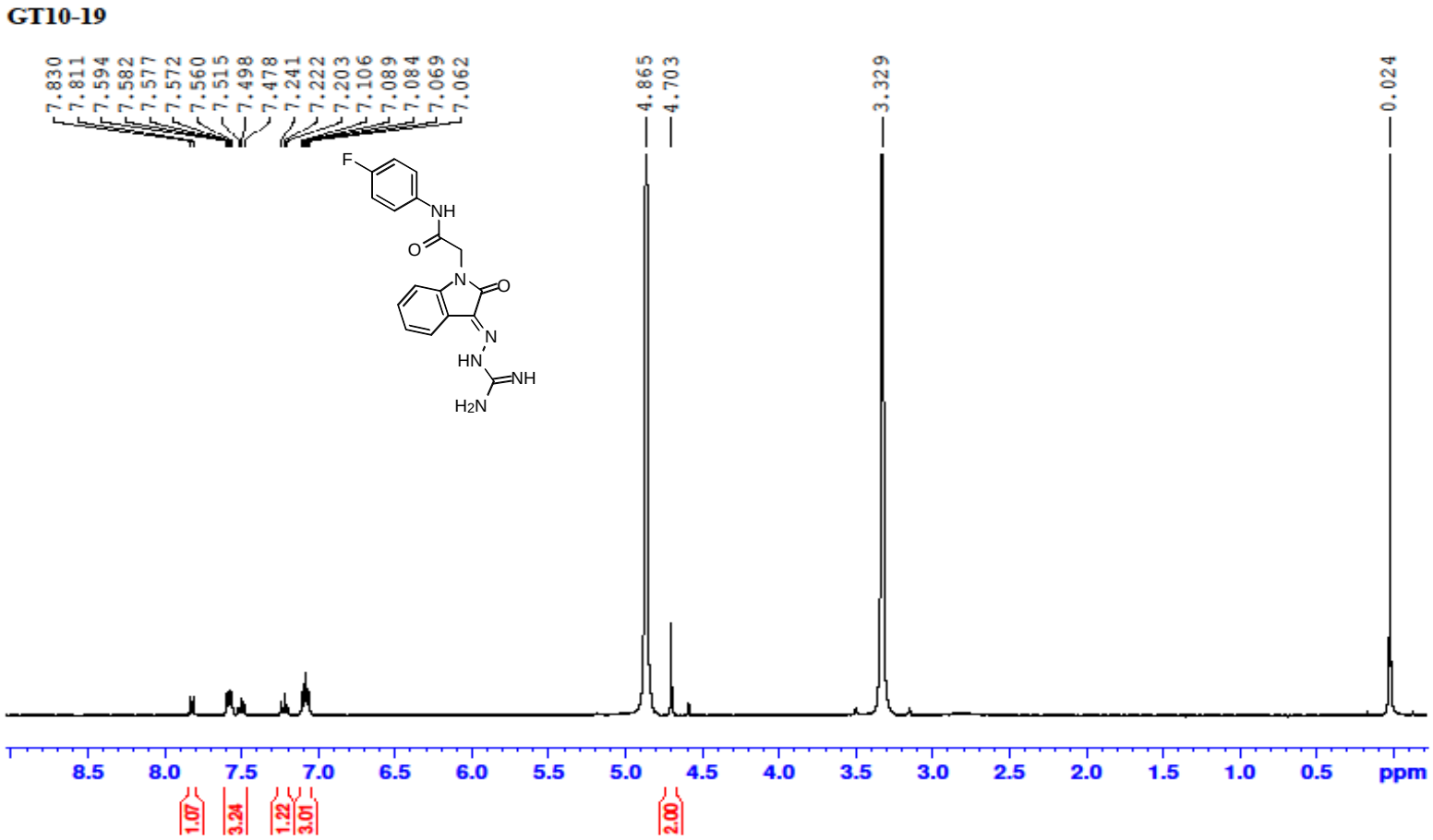
I-17 ^{1}H NMR spectrum of compound GT10-16



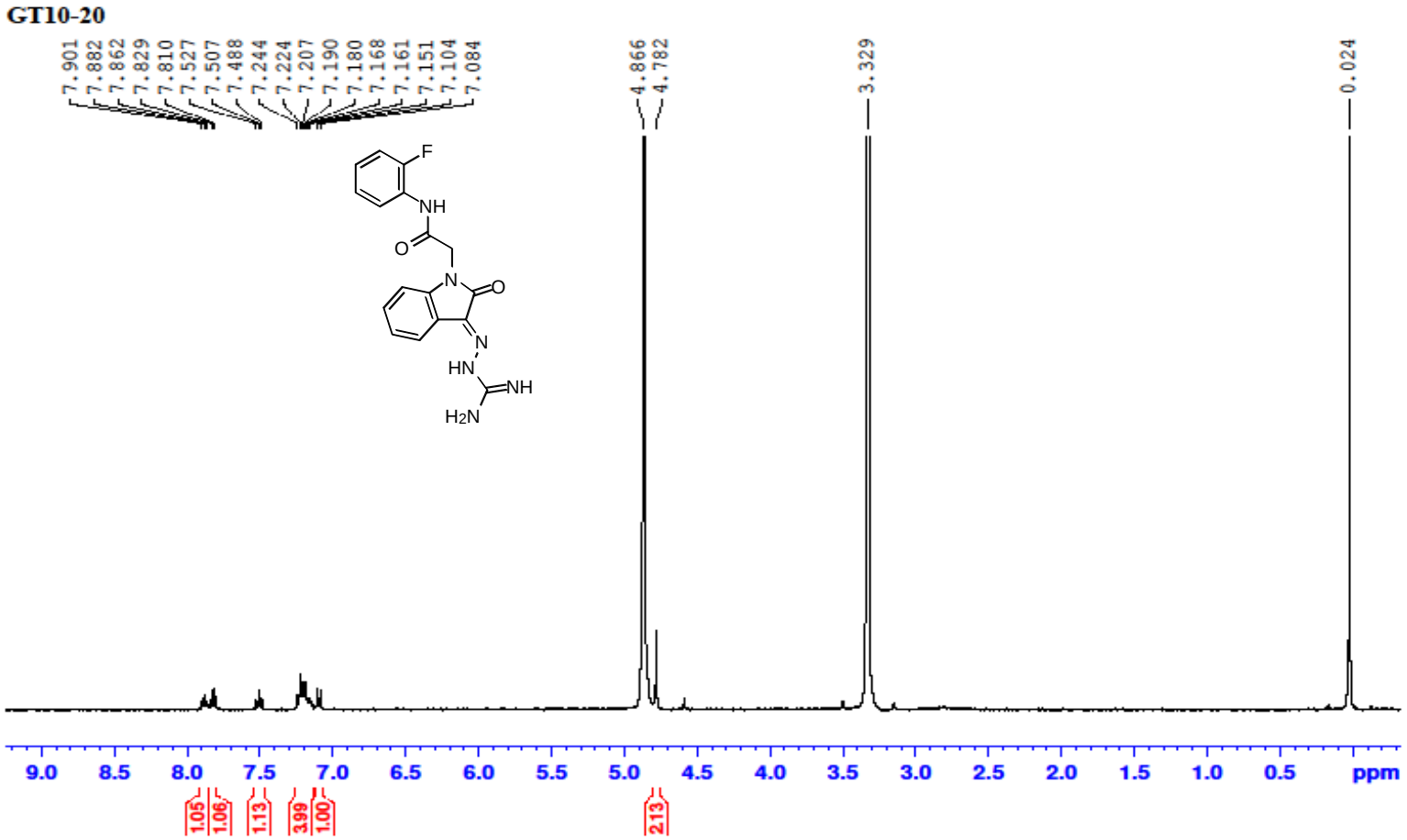
I-18 ^{1}H NMR spectrum of compound GT10-17



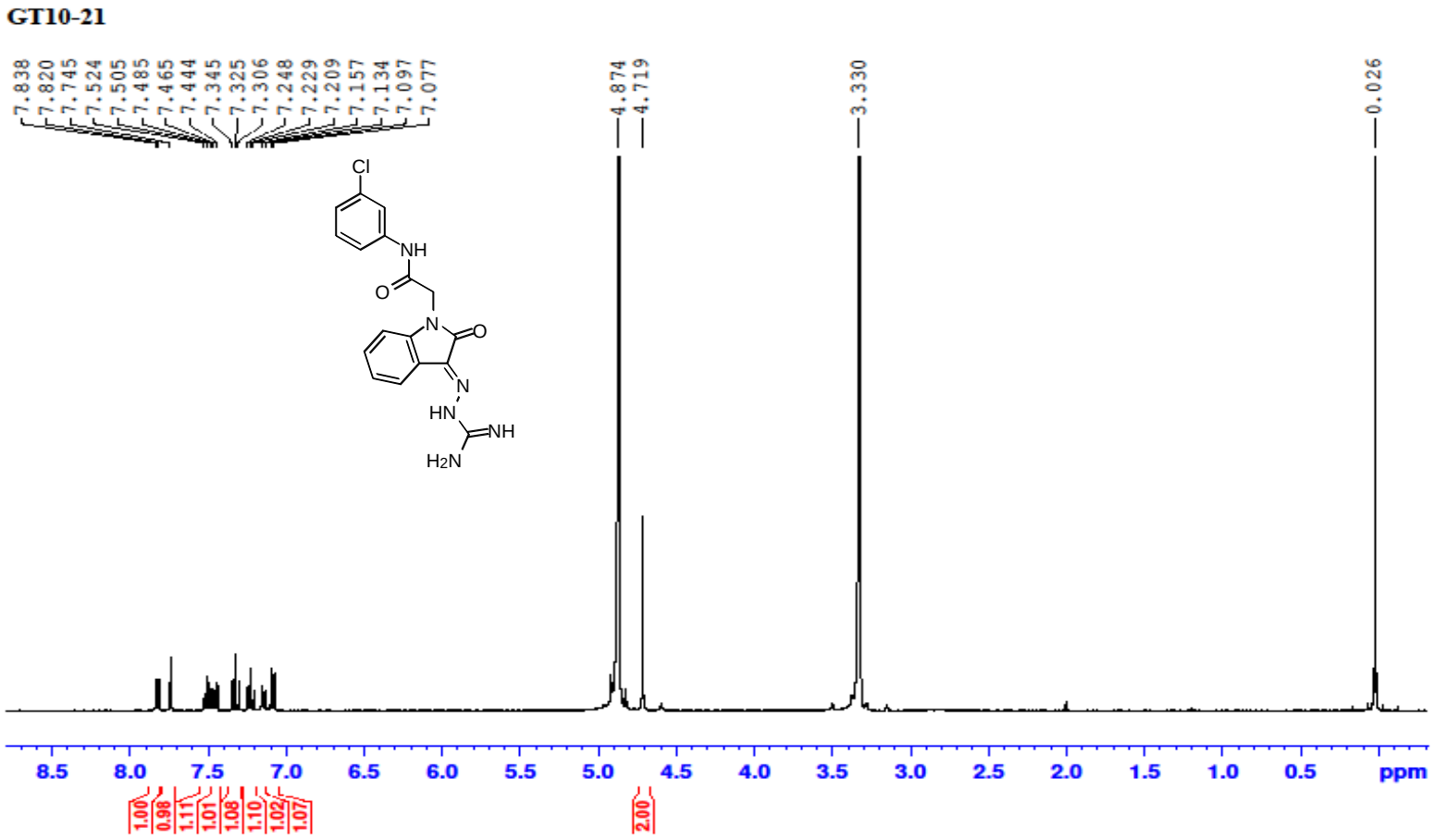
I-19 ^{1}H NMR spectrum of compound GT10-18



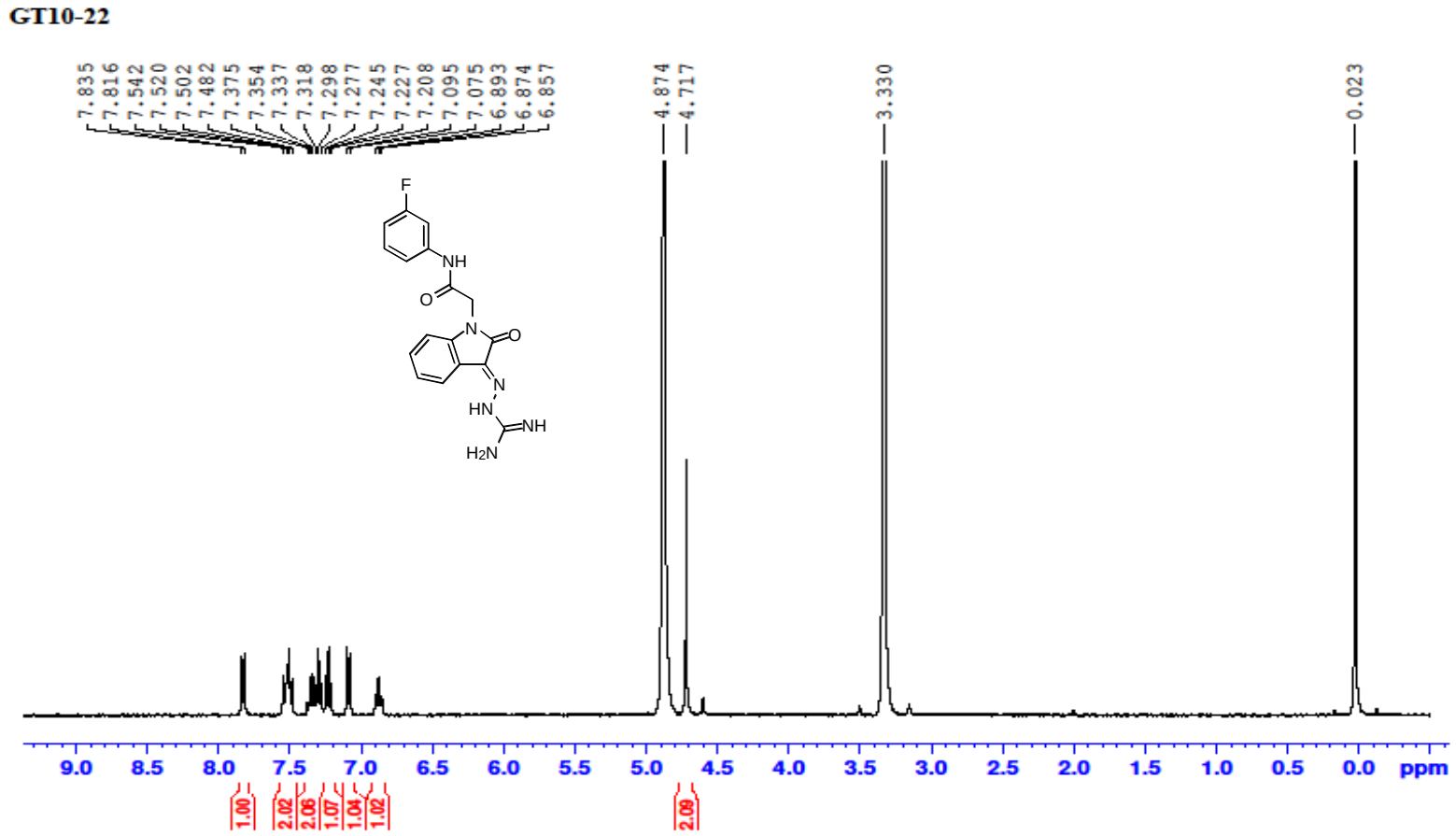
I-20 ^{1}H NMR spectrum of compound GT10-19



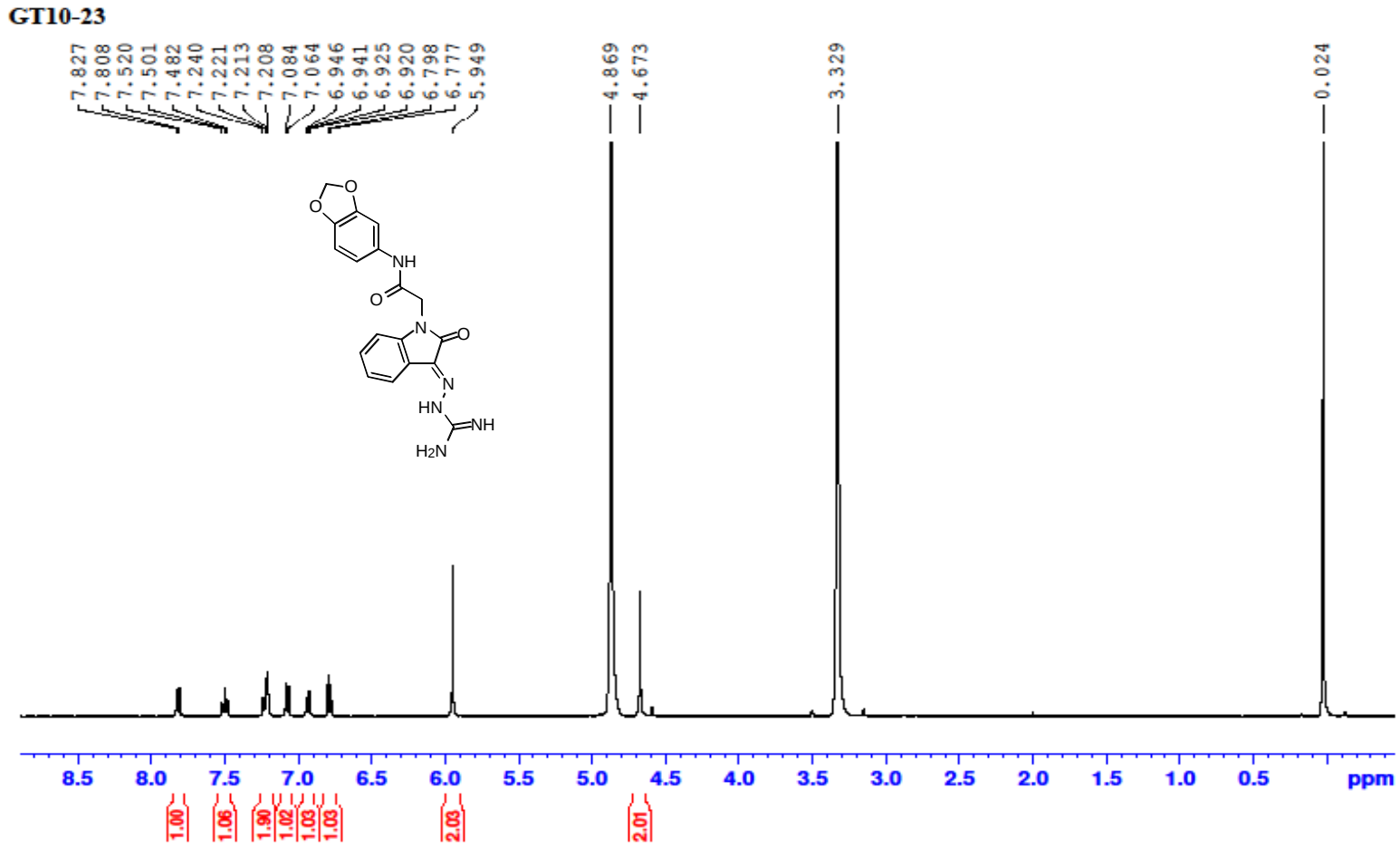
I-21 ^{1}H NMR spectrum of compound GT10-20



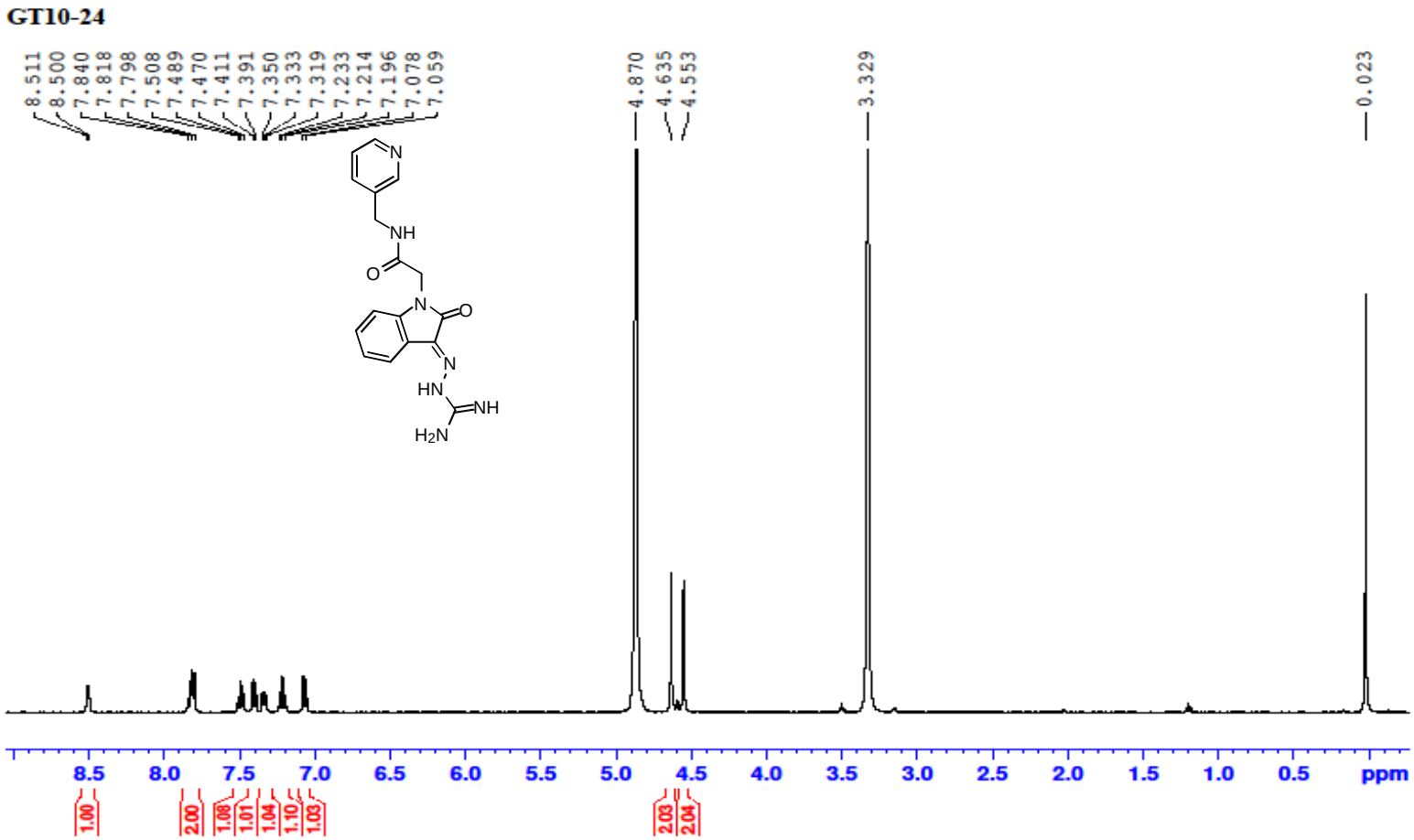
I-22 ^{1}H NMR spectrum of compound GT10-21



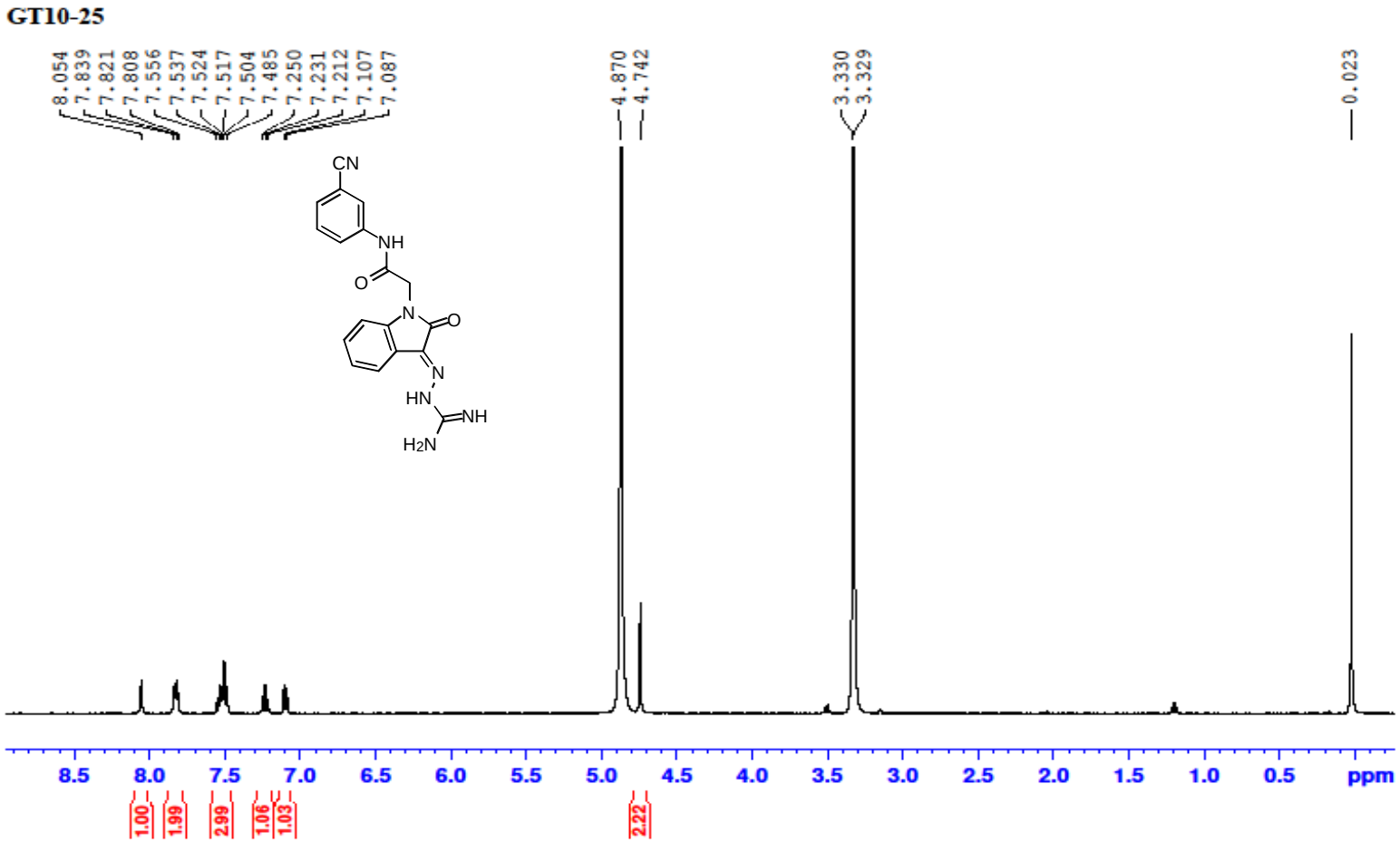
I-23 ^{1}H NMR spectrum of compound GT10-22



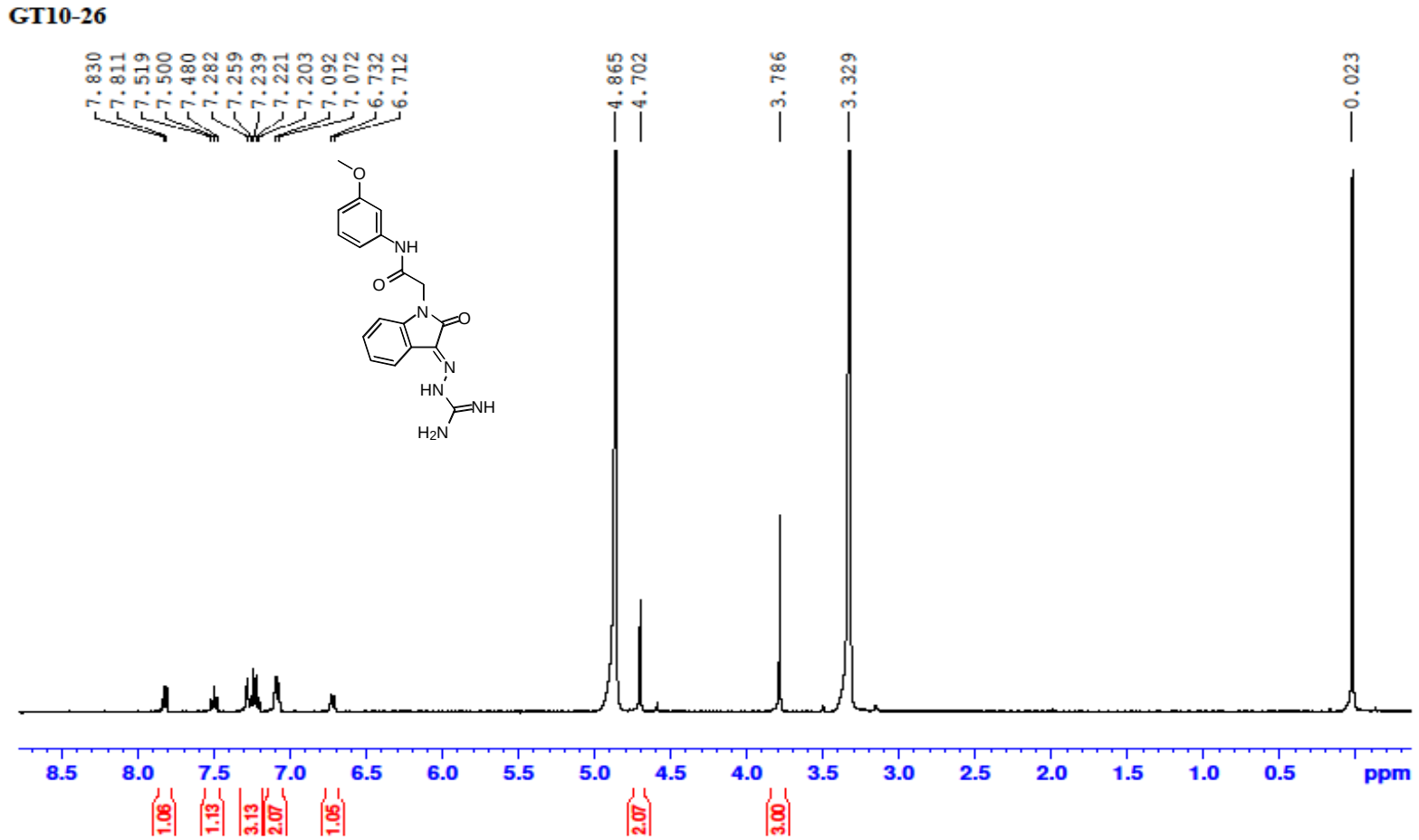
I-24 ^{1}H NMR spectrum of compound GT10-23



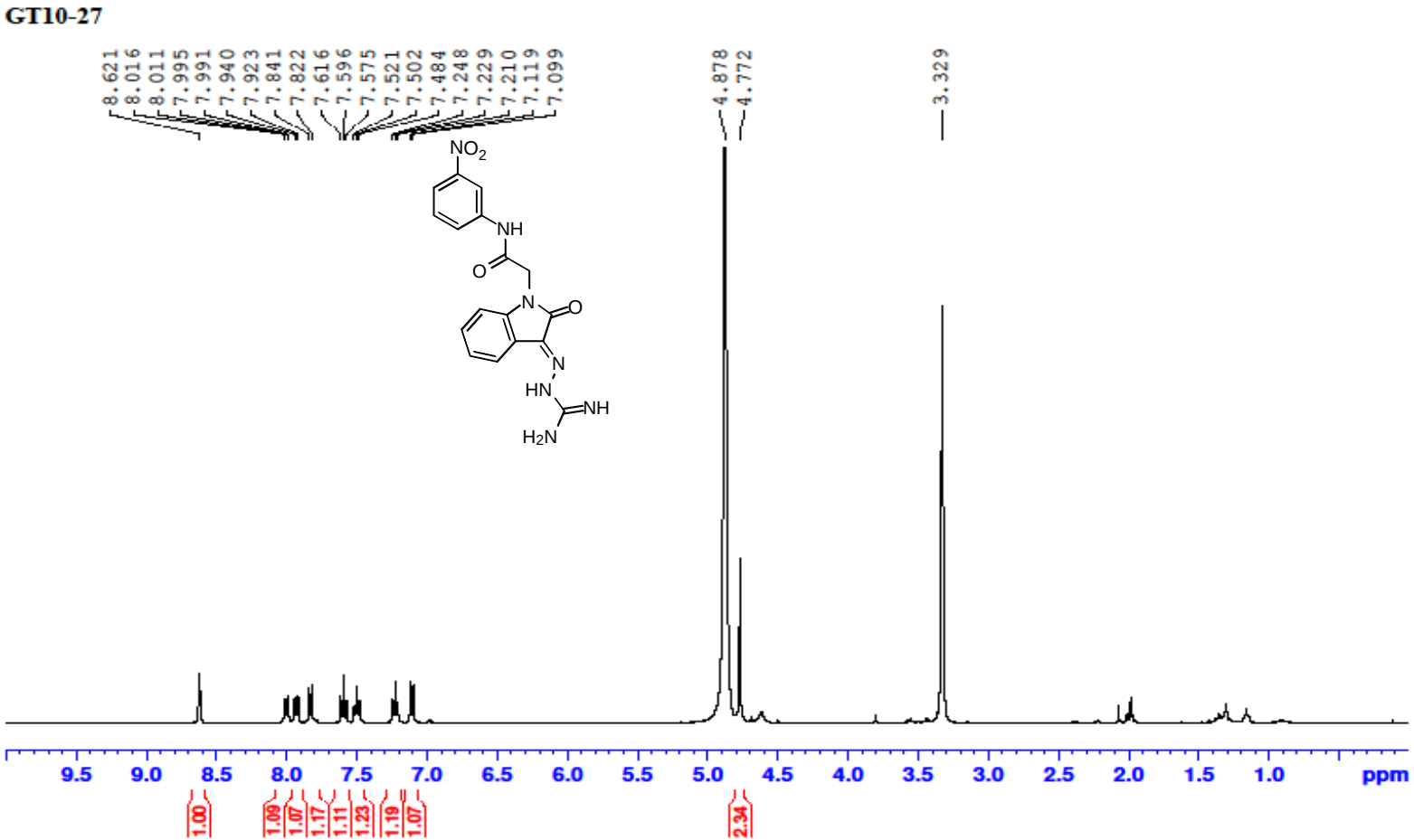
I-25 ^{1}H NMR spectrum of compound GT10-24



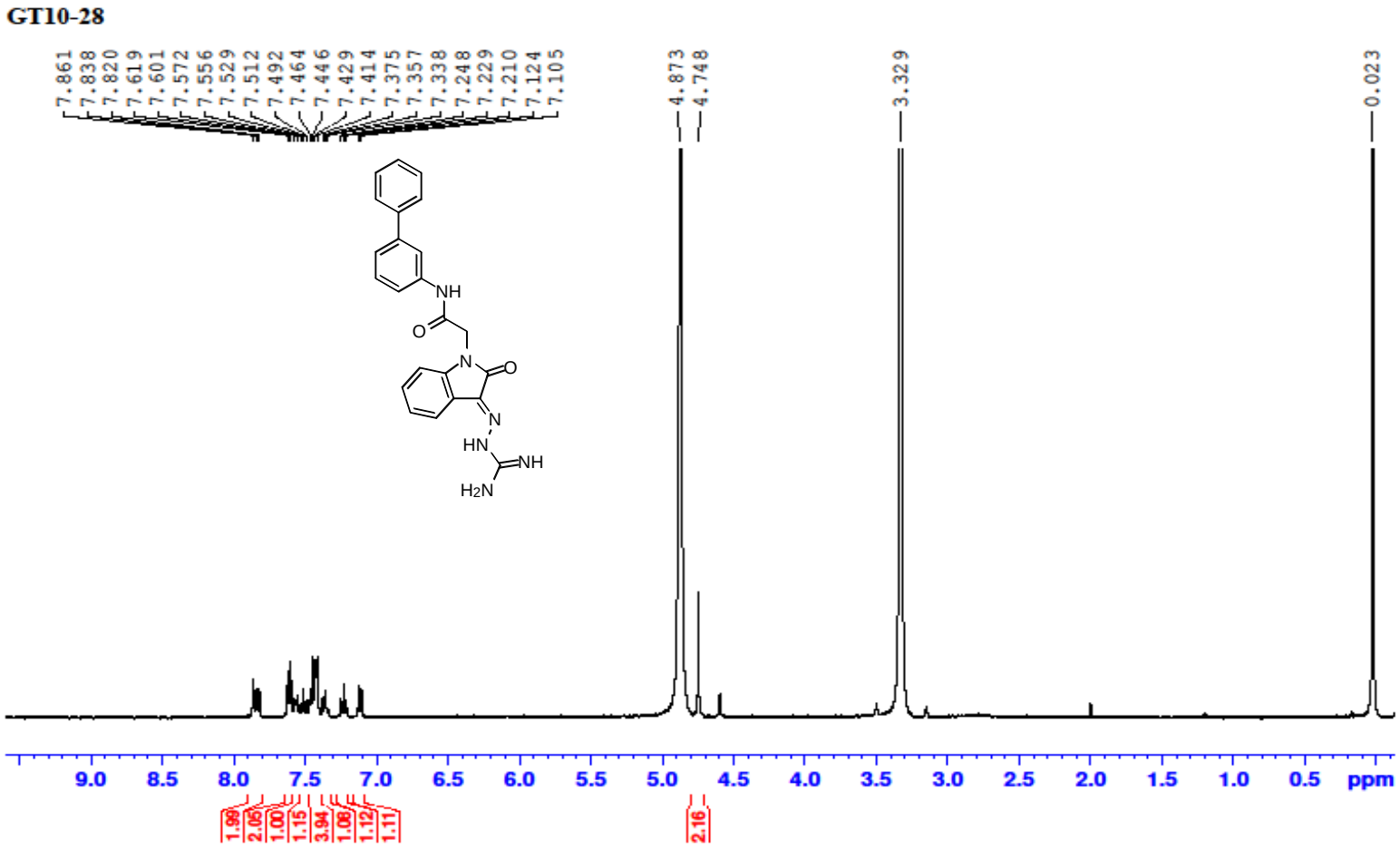
I-26 ^{1}H NMR spectrum of compound GT10-25



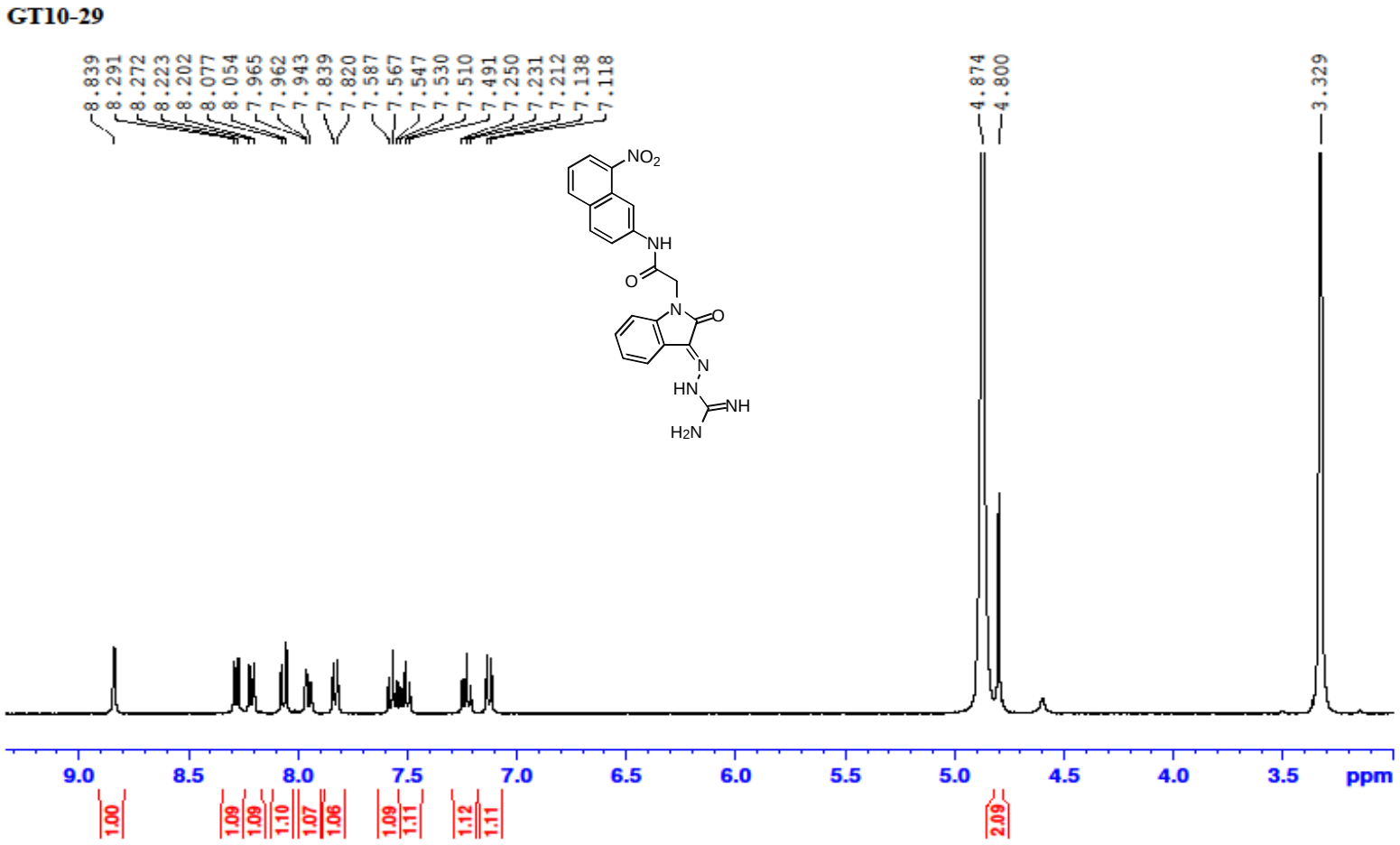
I-27 ^{1}H NMR spectrum of compound GT10-26



I-28 ^{1}H NMR spectrum of compound GT10-27



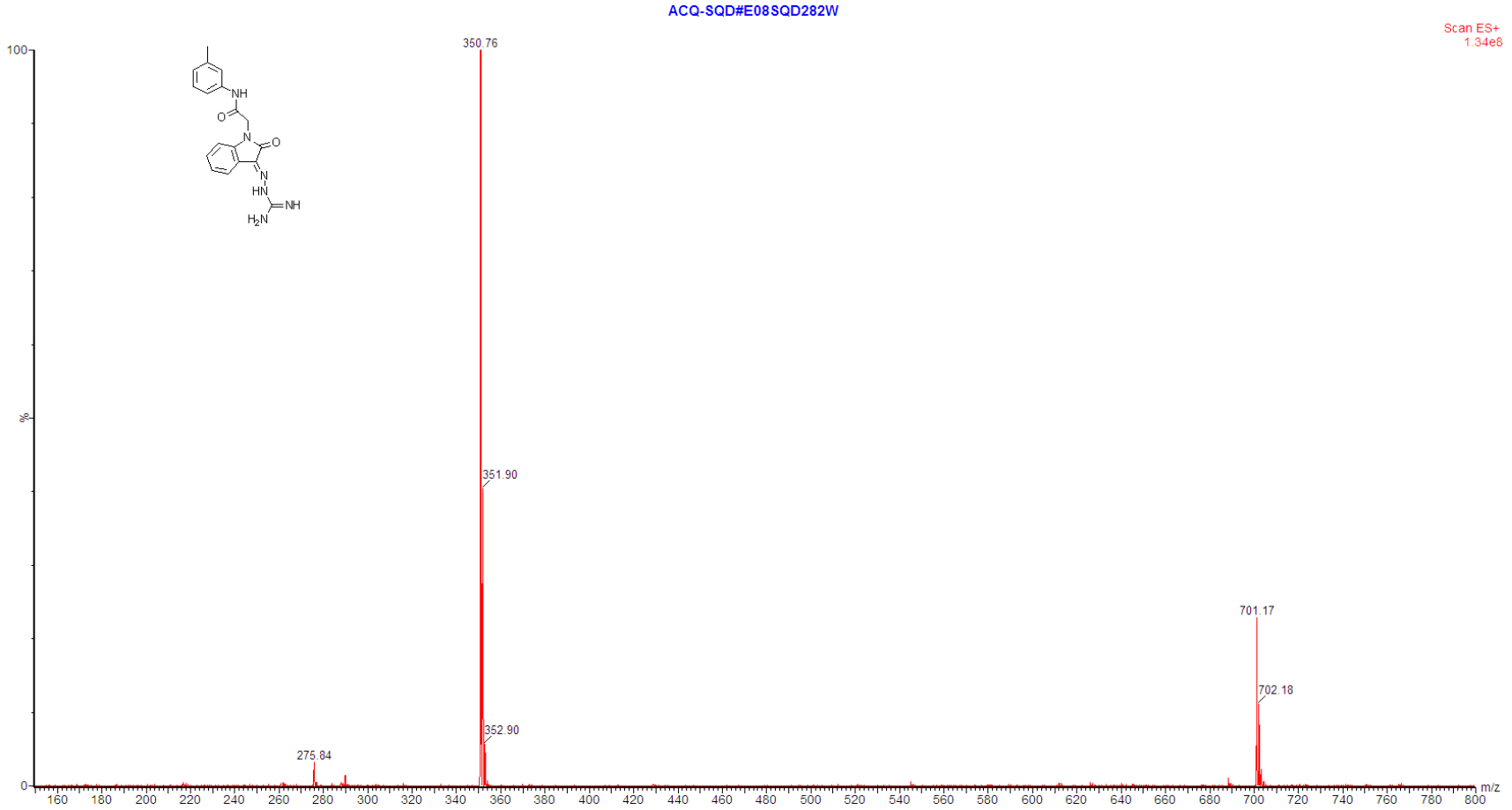
I-29 ^{1}H NMR spectrum of compound GT10-28



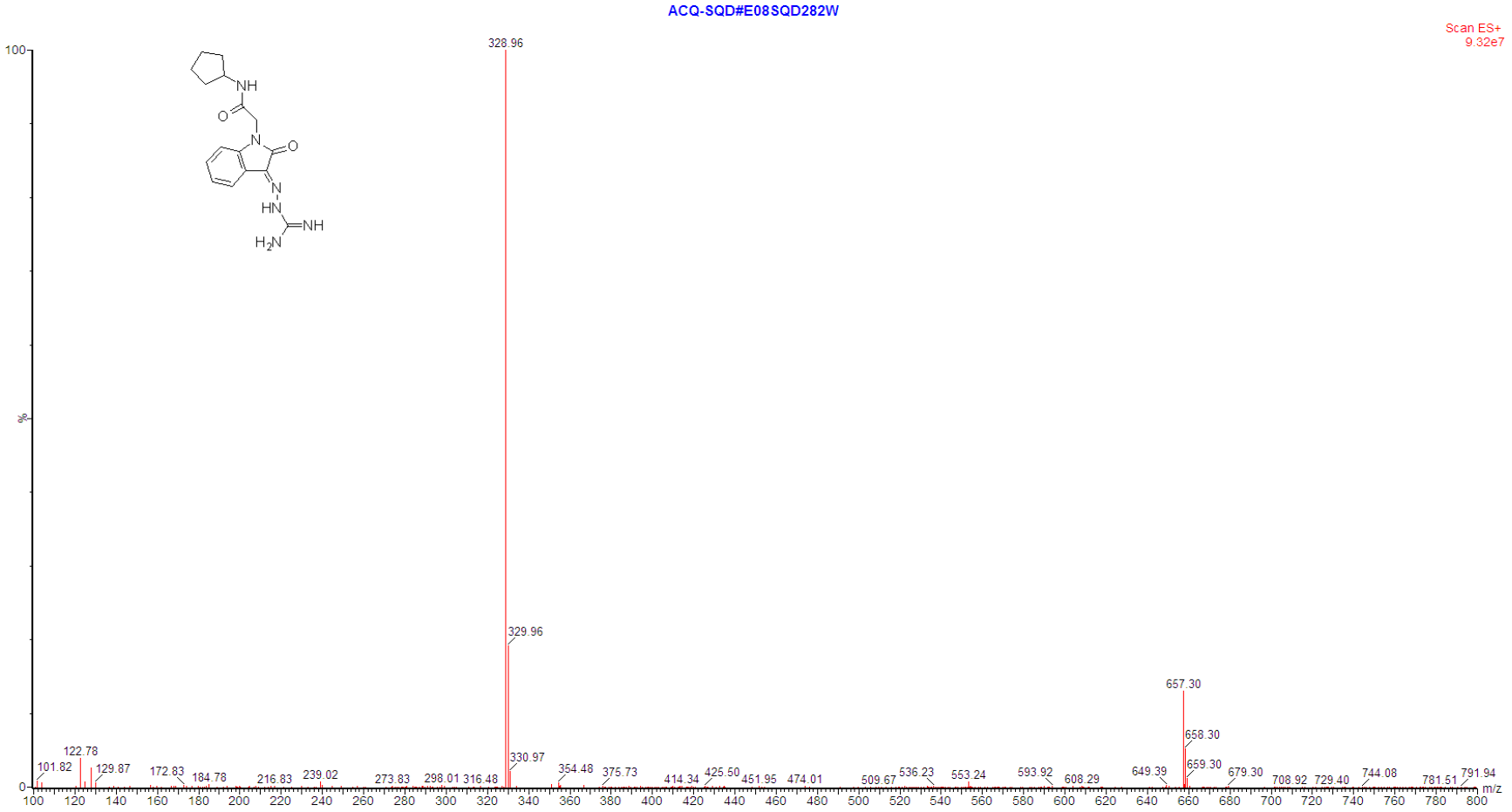
I-30 ^{1}H NMR spectrum of compound GT10-29

Appendix II

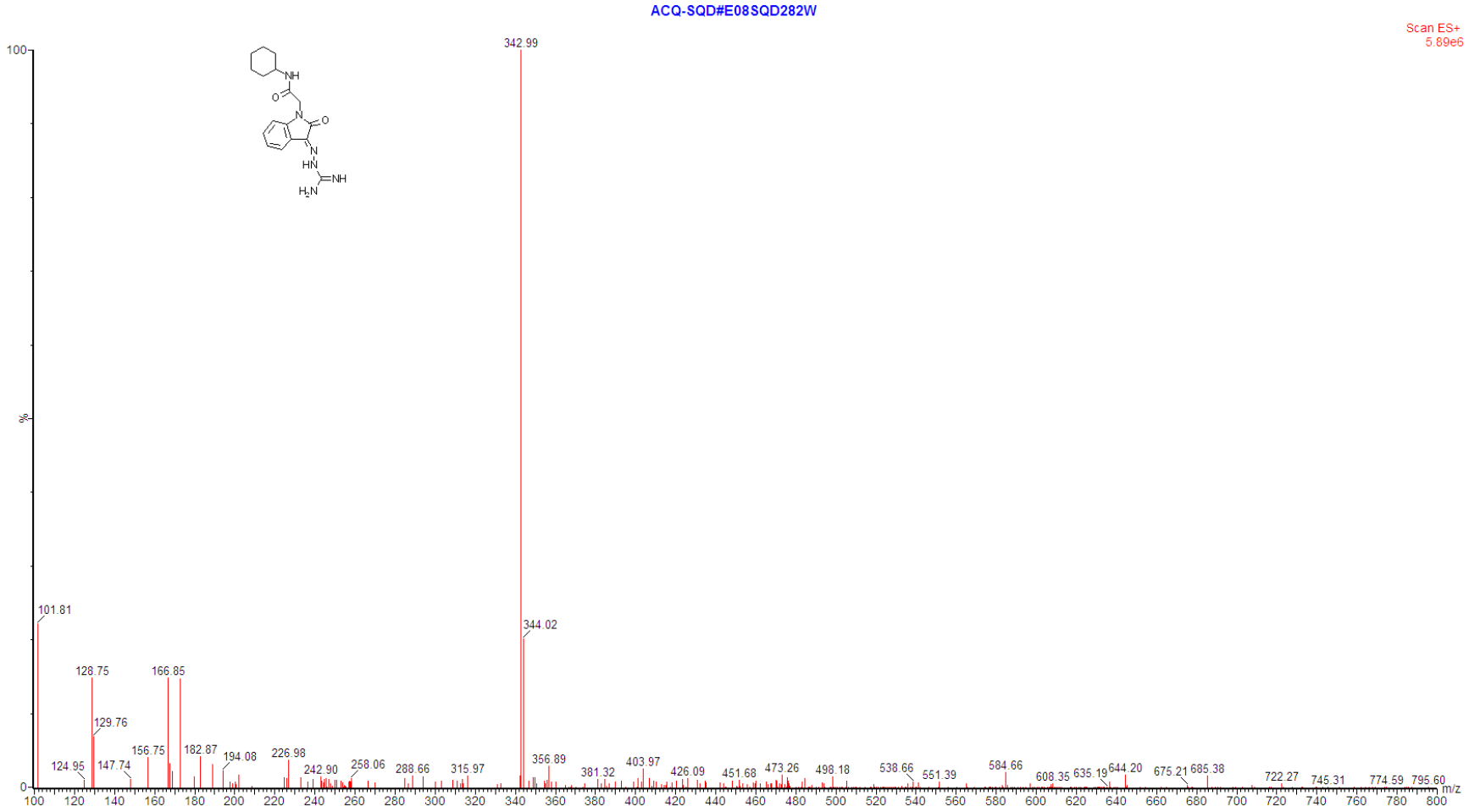
MS spectra of GT10 and its analogues



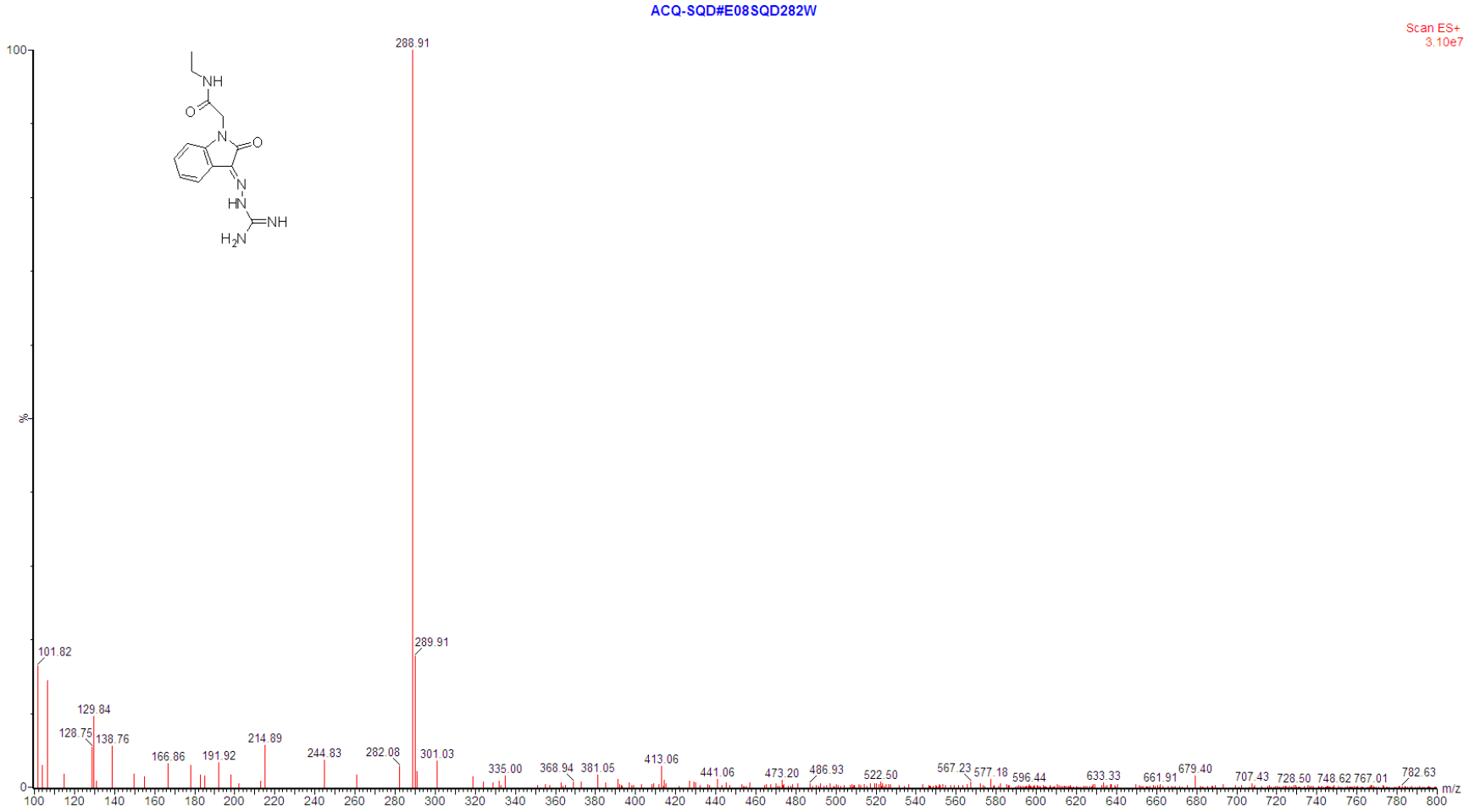
II-1 MS spectrum of compound GT10



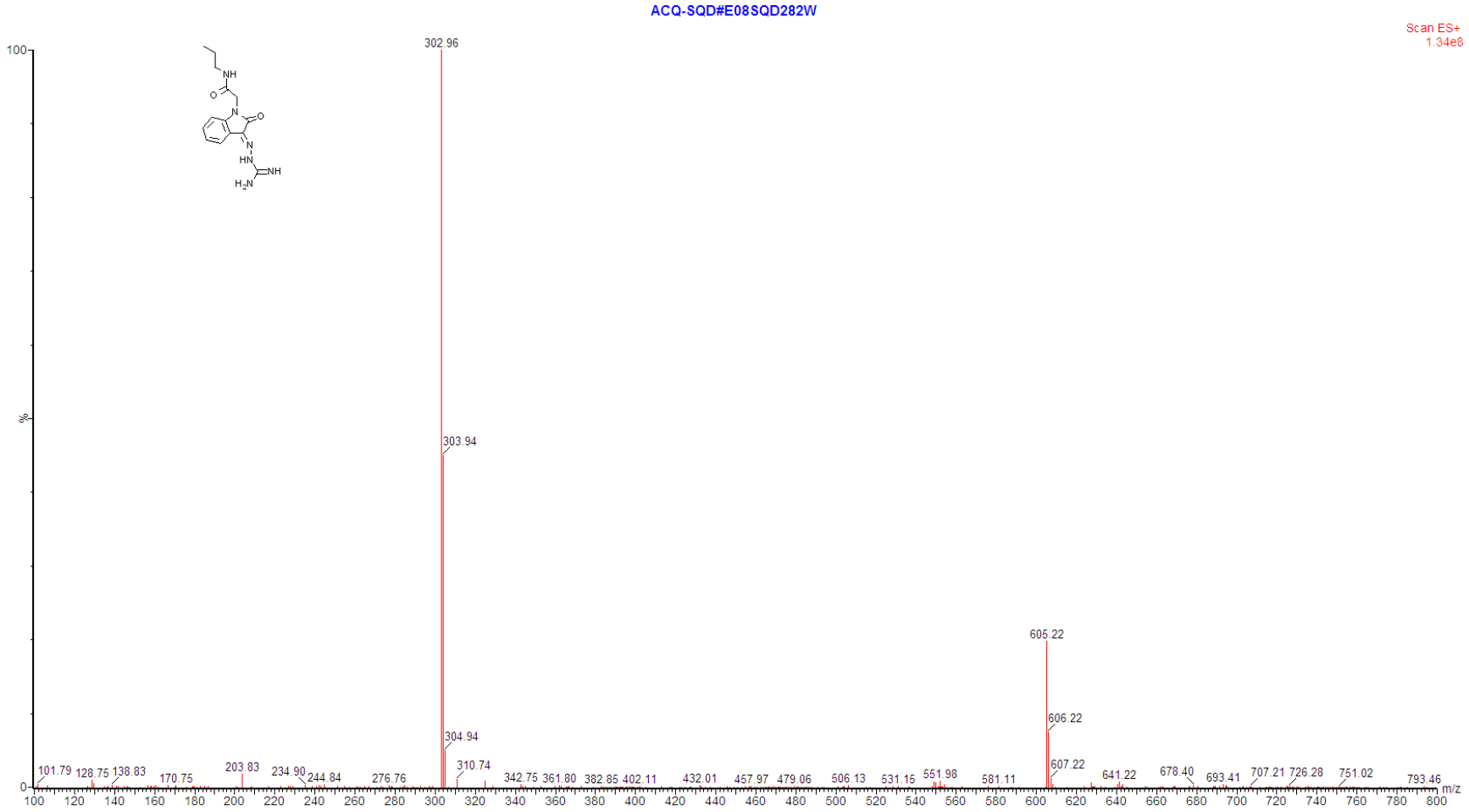
II-2 MS spectrum of compound GT10-1



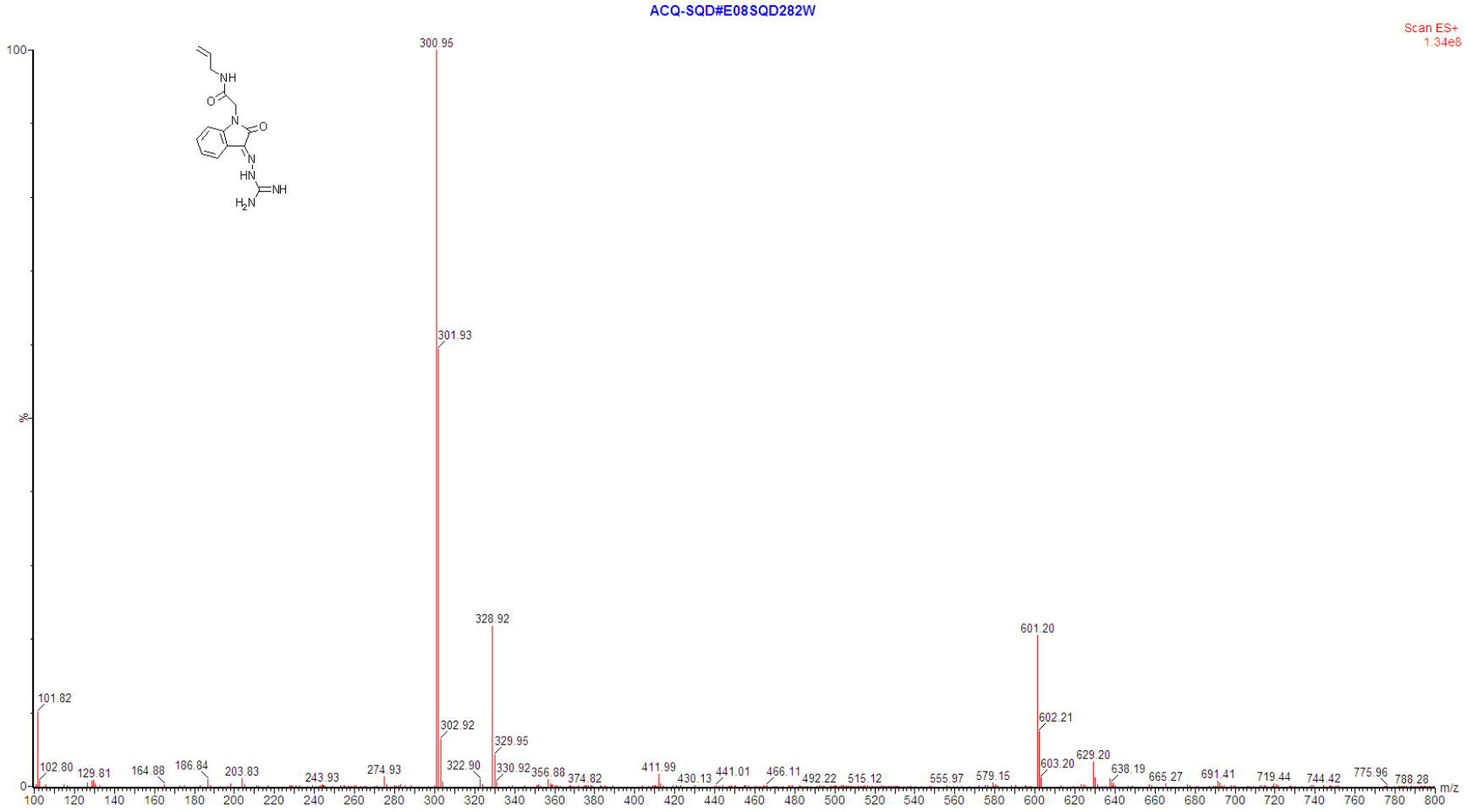
II-3 MS spectrum of compound GT10-2



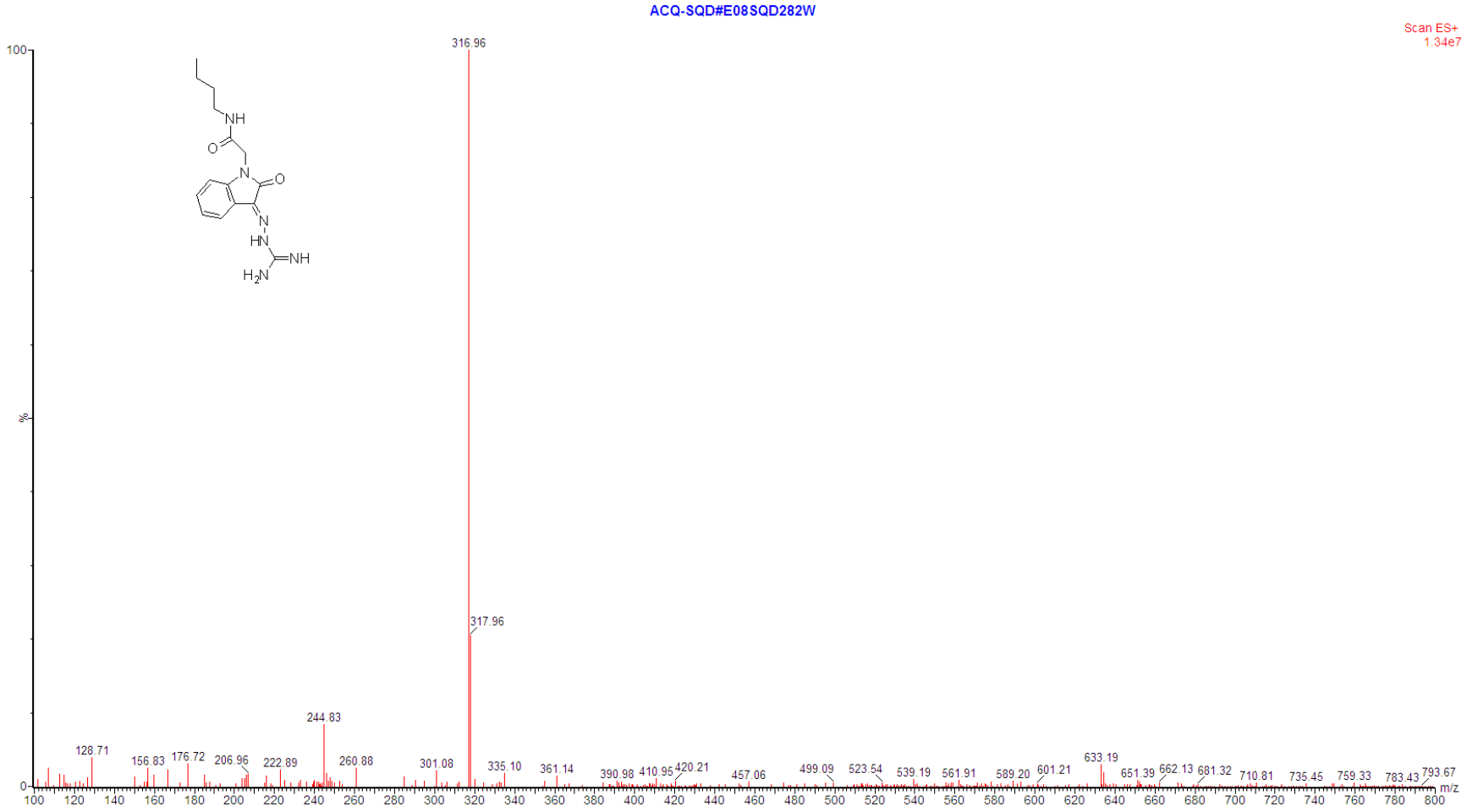
II-4 MS spectrum of compound GT10-3



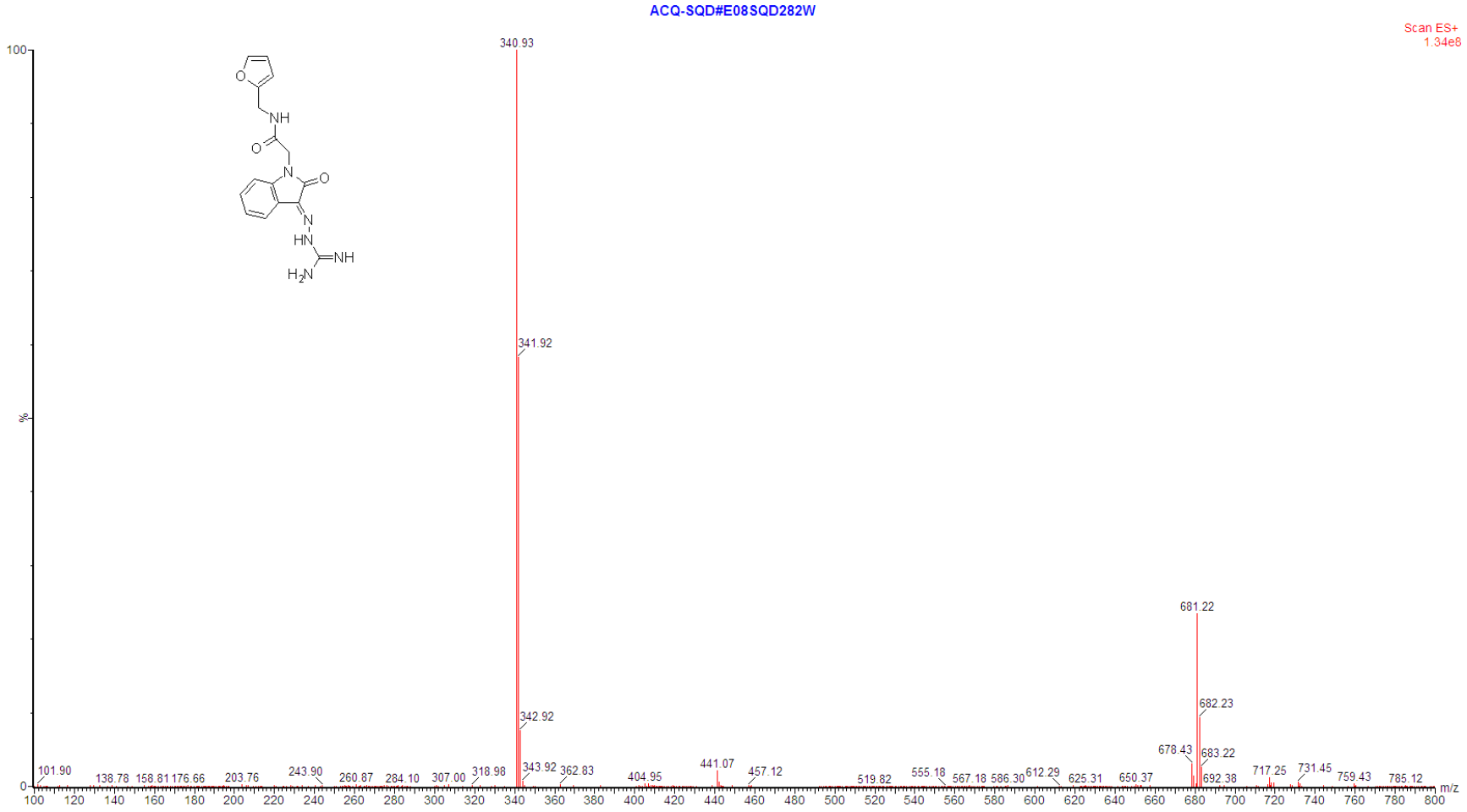
II-5 MS spectrum of compound GT10-4



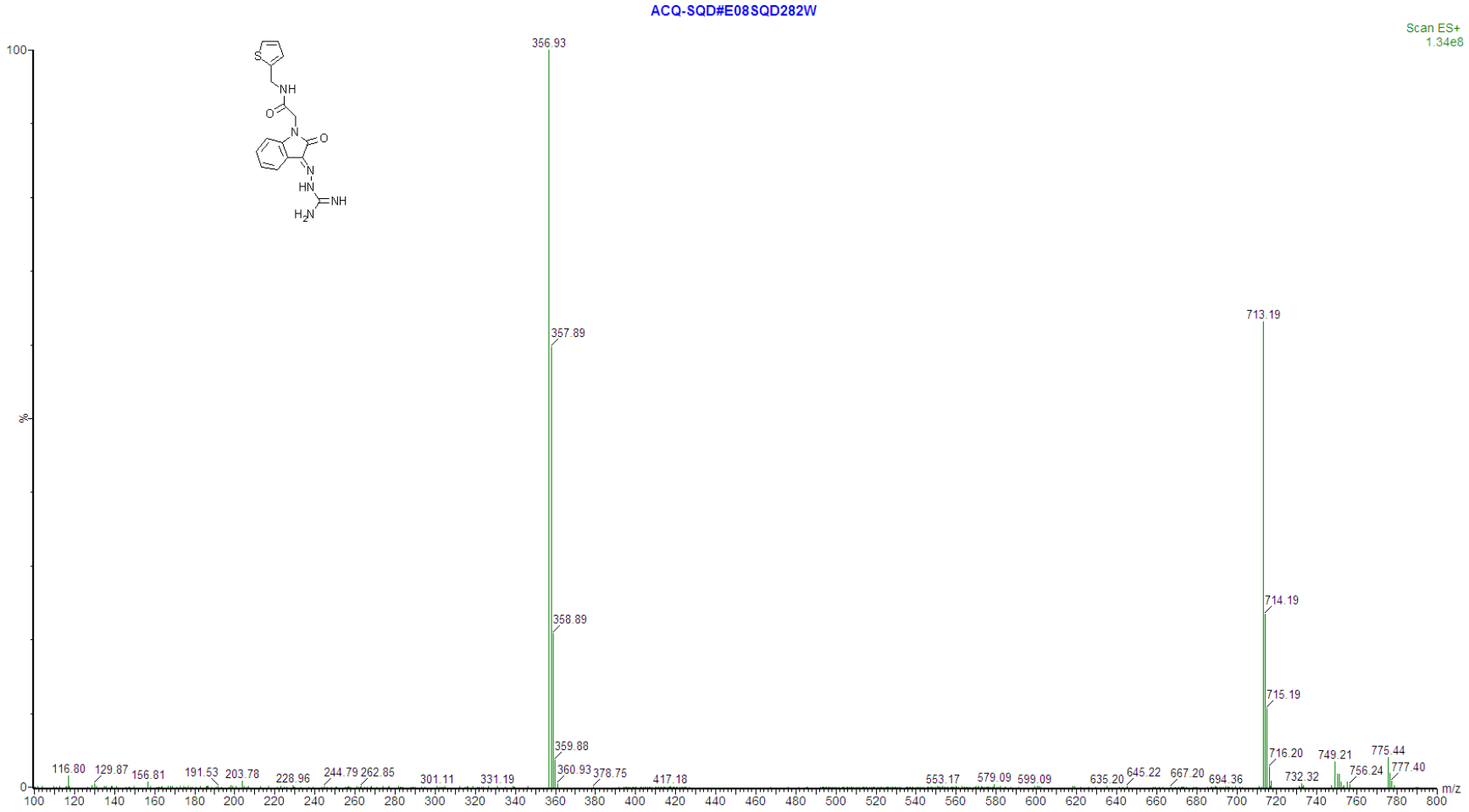
II-6 MS spectrum of compound GT10-5



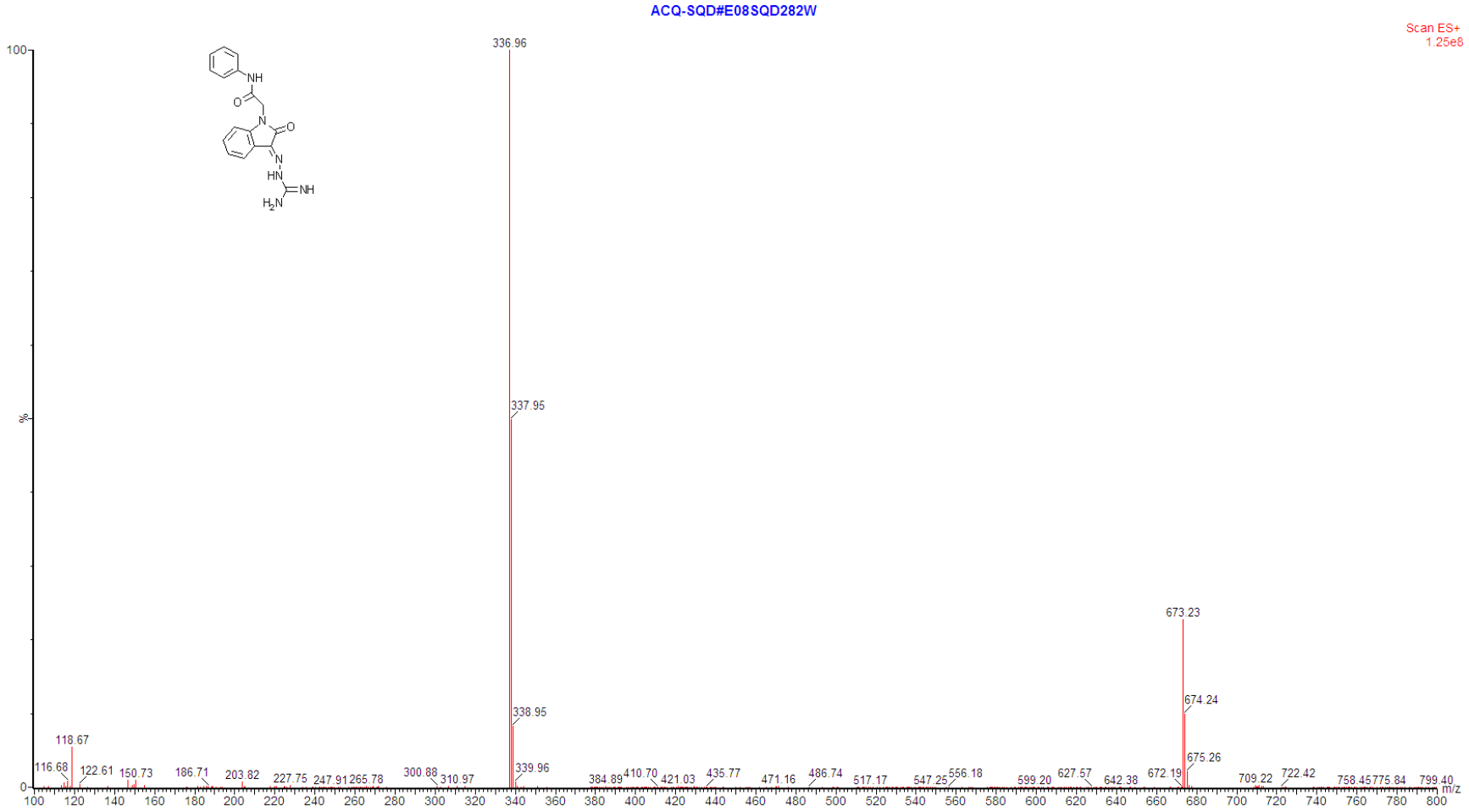
II-7 MS spectrum of compound GT10-6



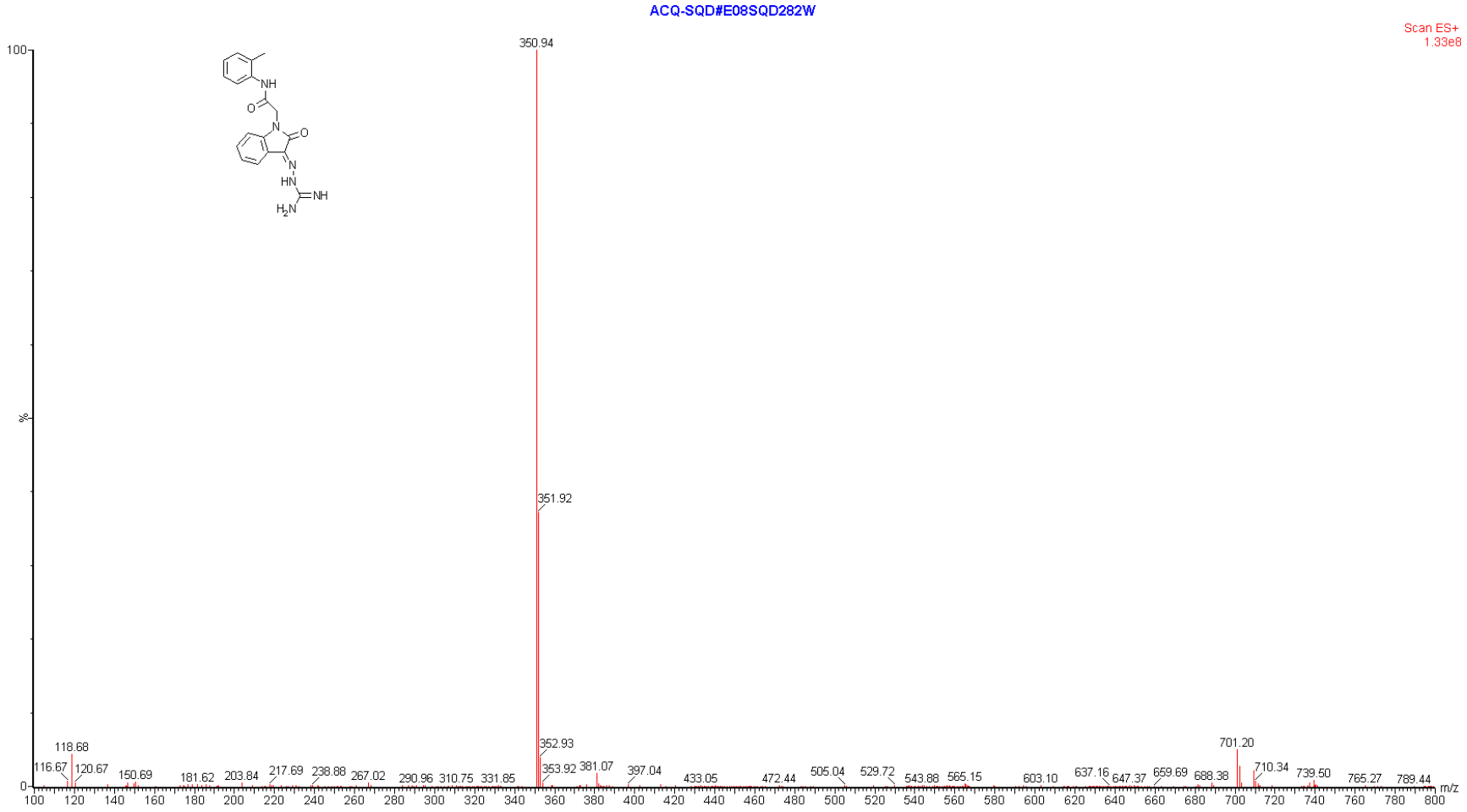
II-8 MS spectrum of compound GT10-7



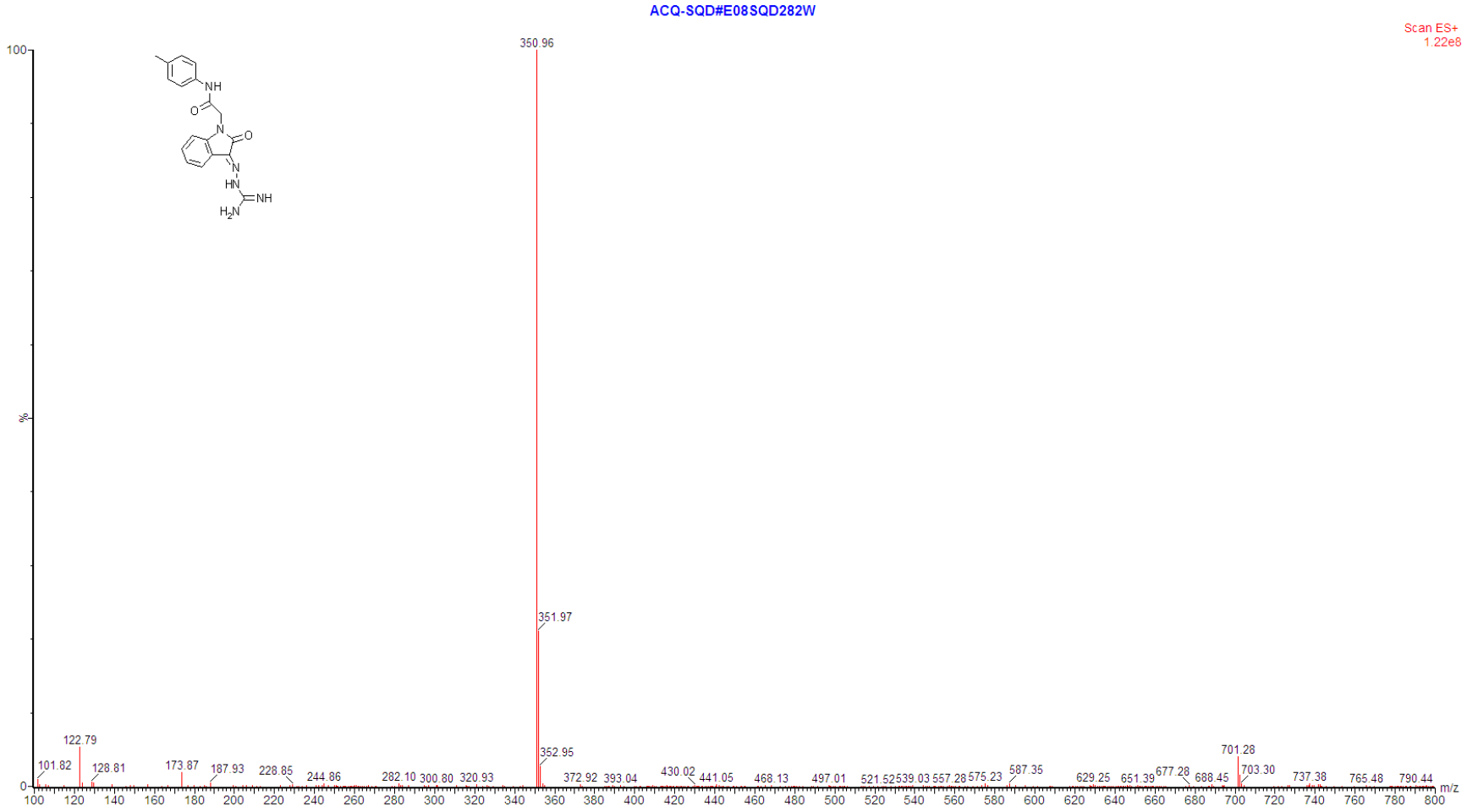
II-9 MS spectrum of compound GT10-8



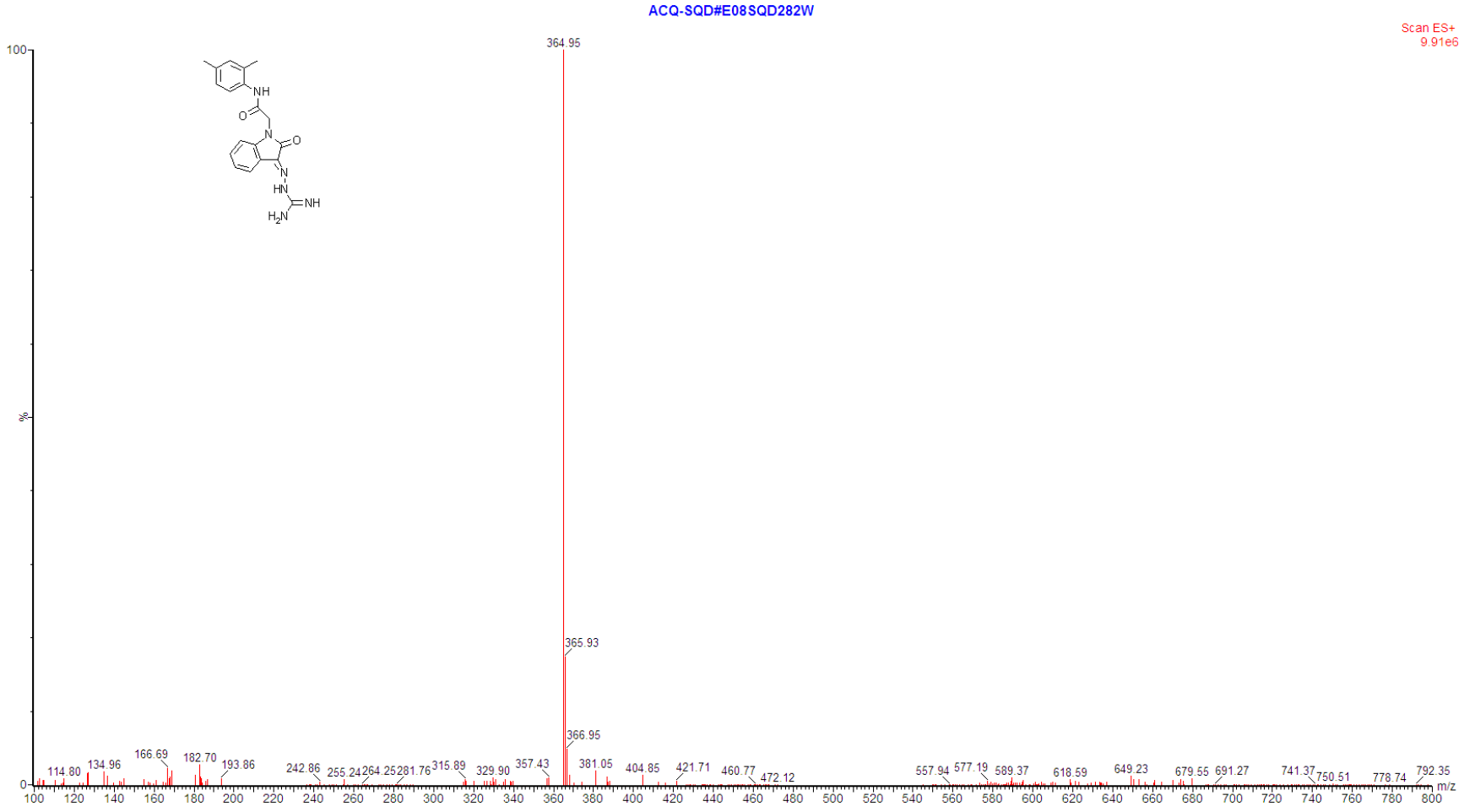
II-10 MS spectrum of compound GT10-9



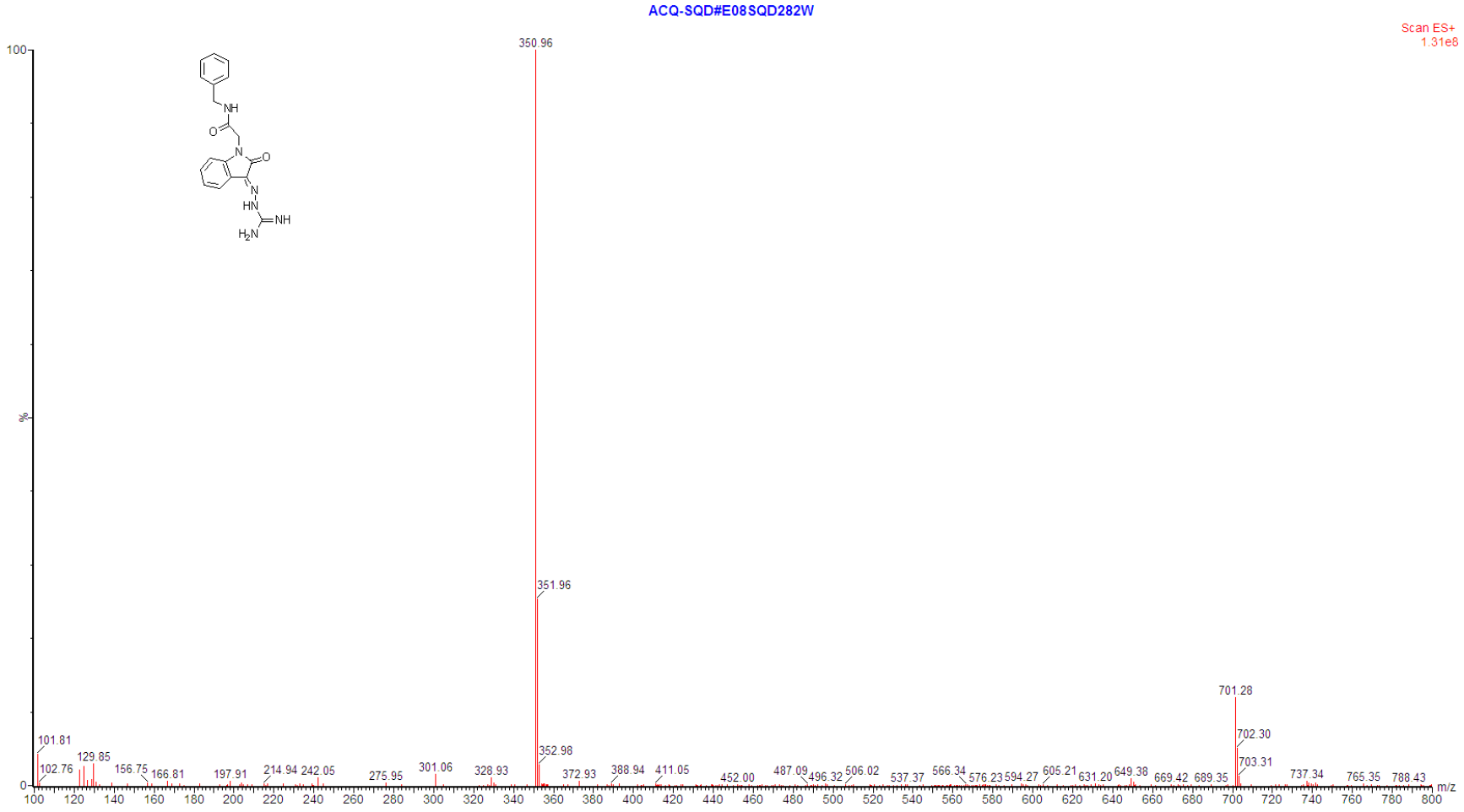
II-11 MS spectrum of compound GT10-10



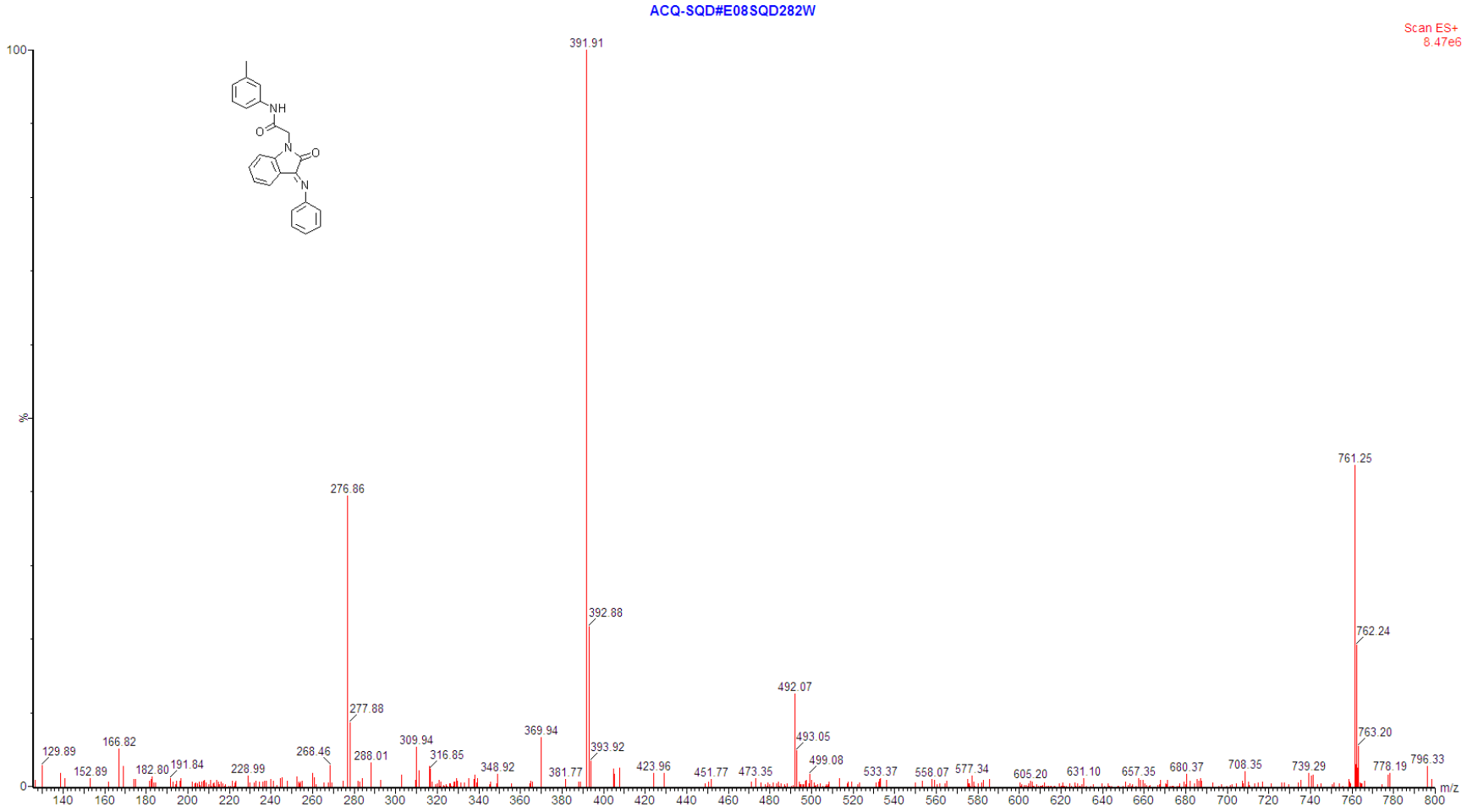
II-12 MS spectrum of compound GT10-11



II-13 MS spectrum of compound GT10-12

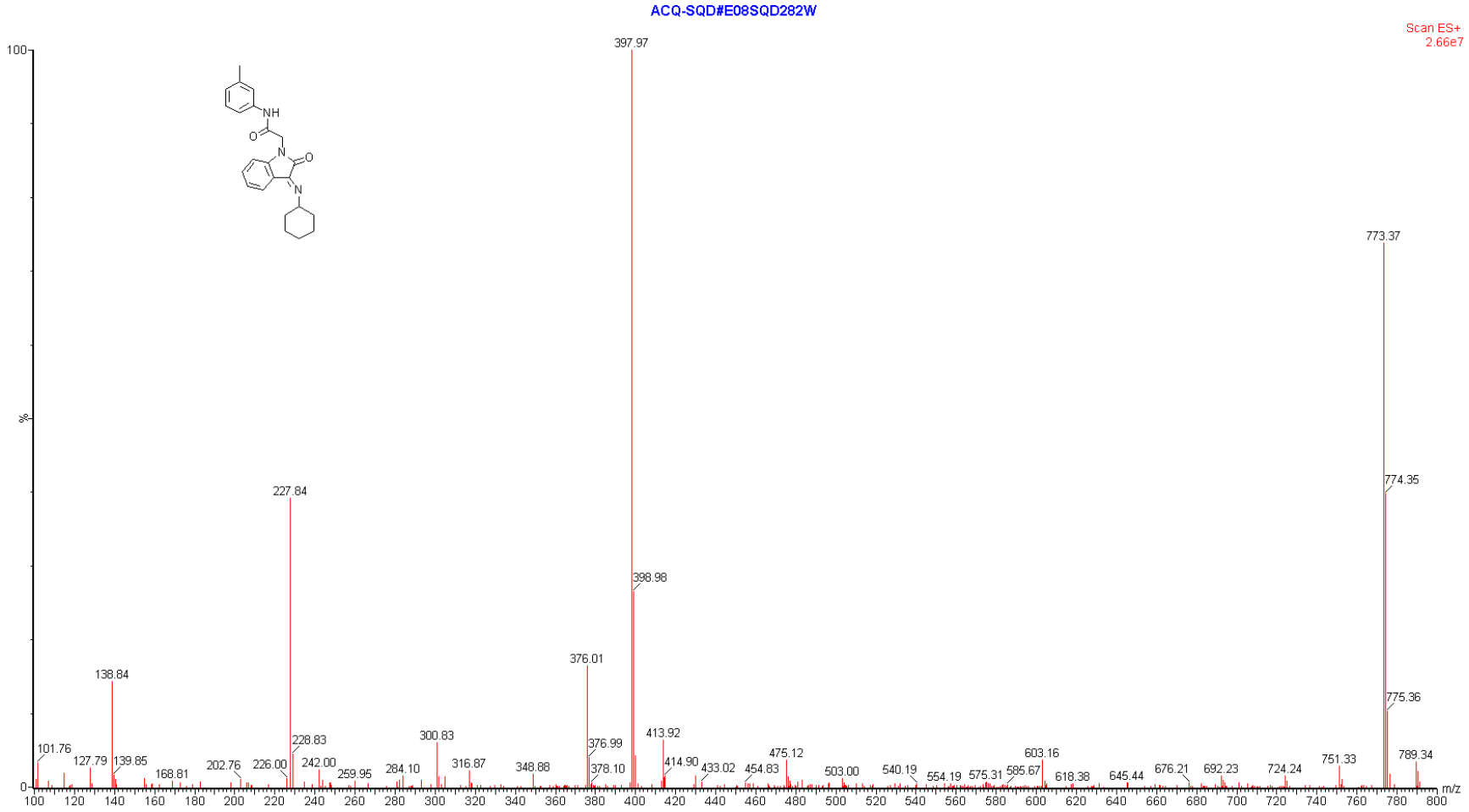


II-14 MS spectrum of compound GT10-13

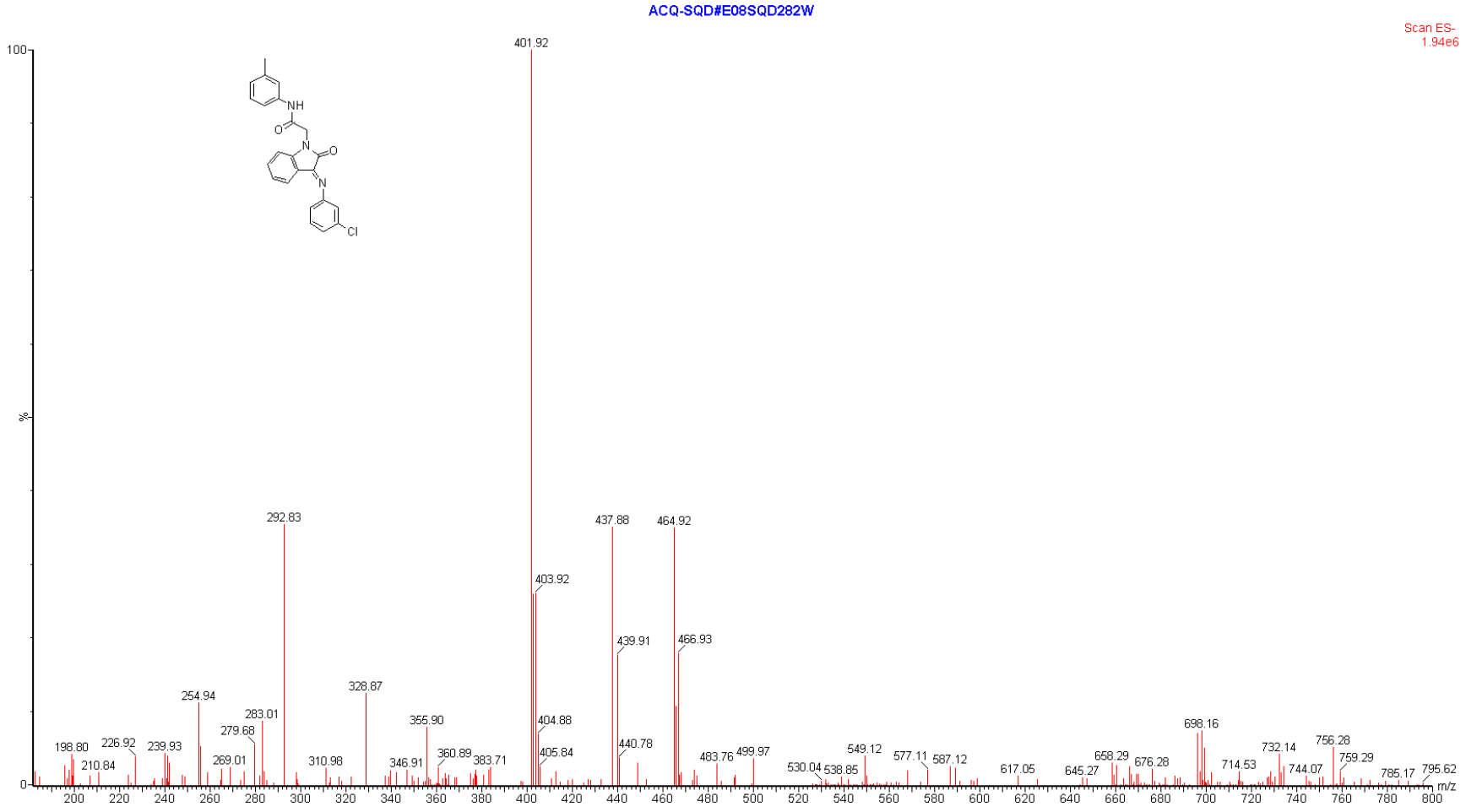


II-15 MS spectrum of compound GT10-14

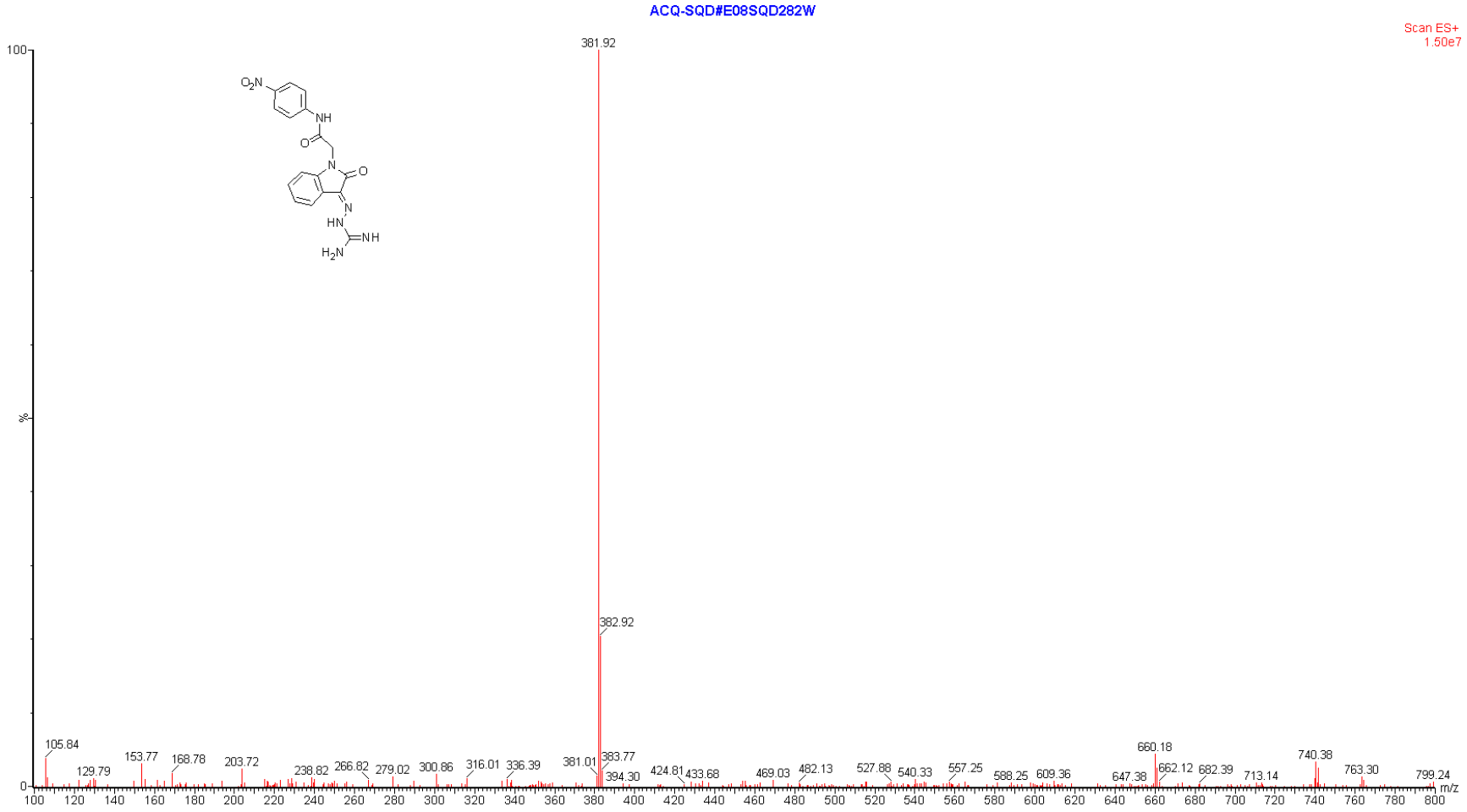
II-16 MS spectrum of compound GT10-15



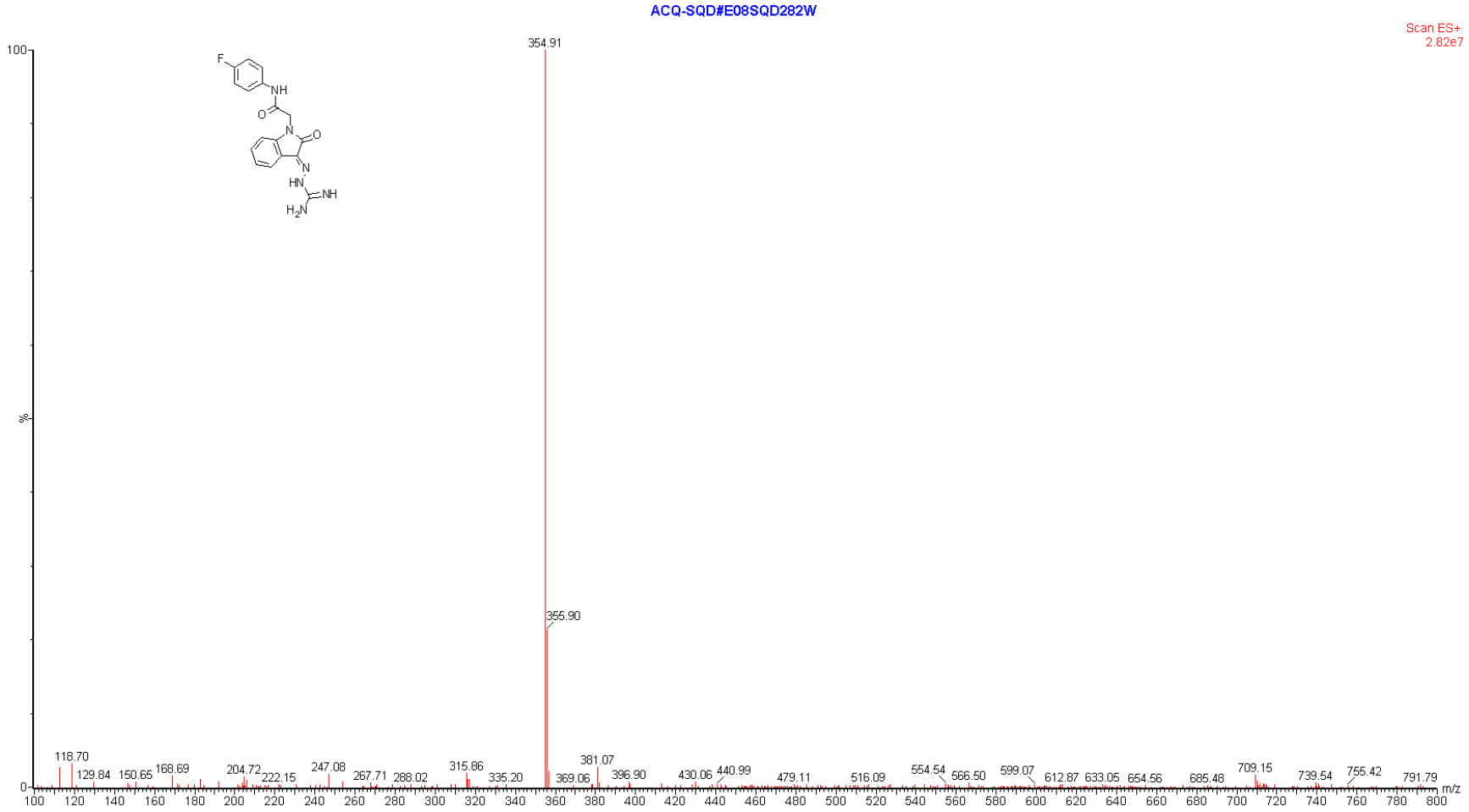
II-17 MS spectrum of compound GT10-16



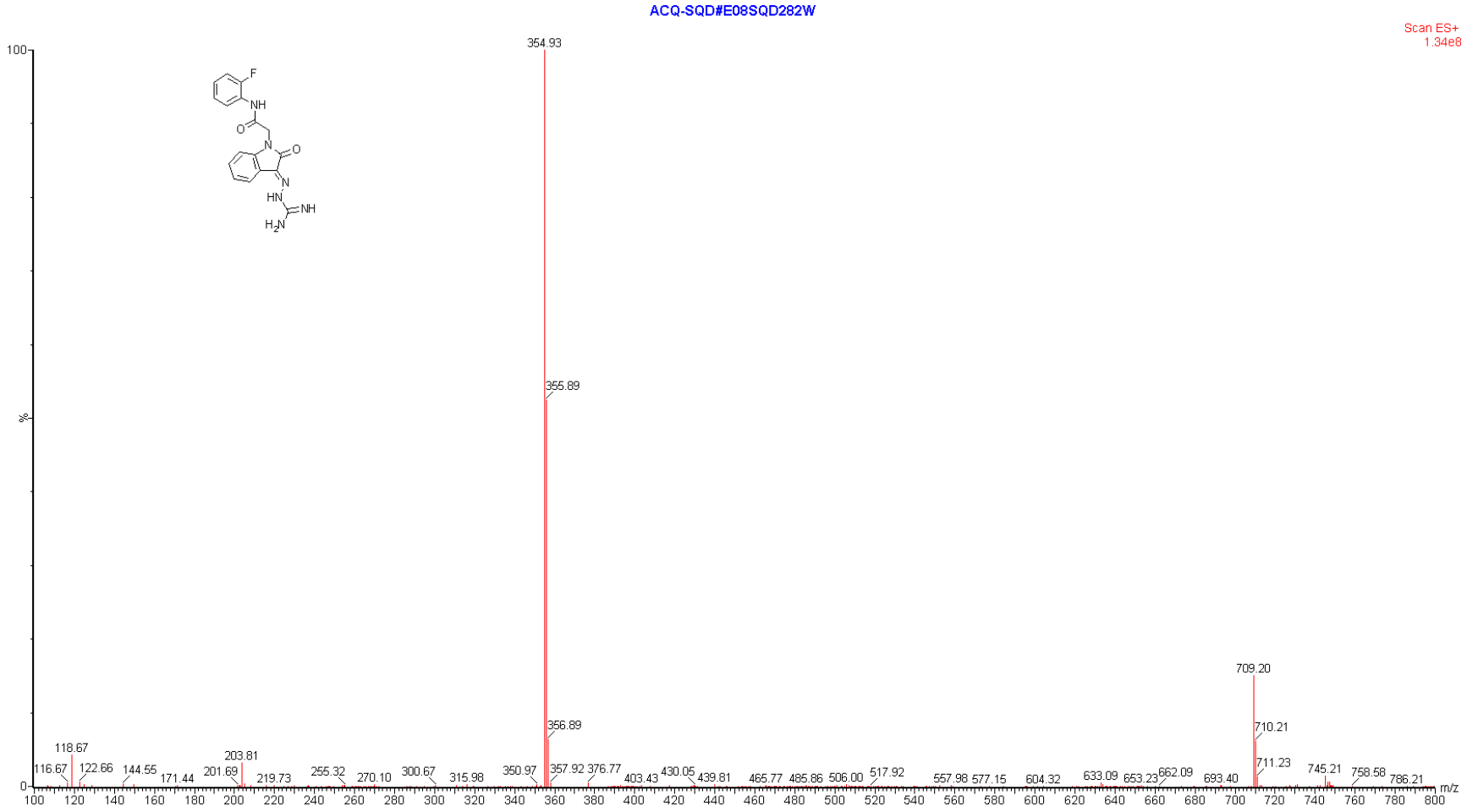
II-18 MS spectrum of compound GT10-17



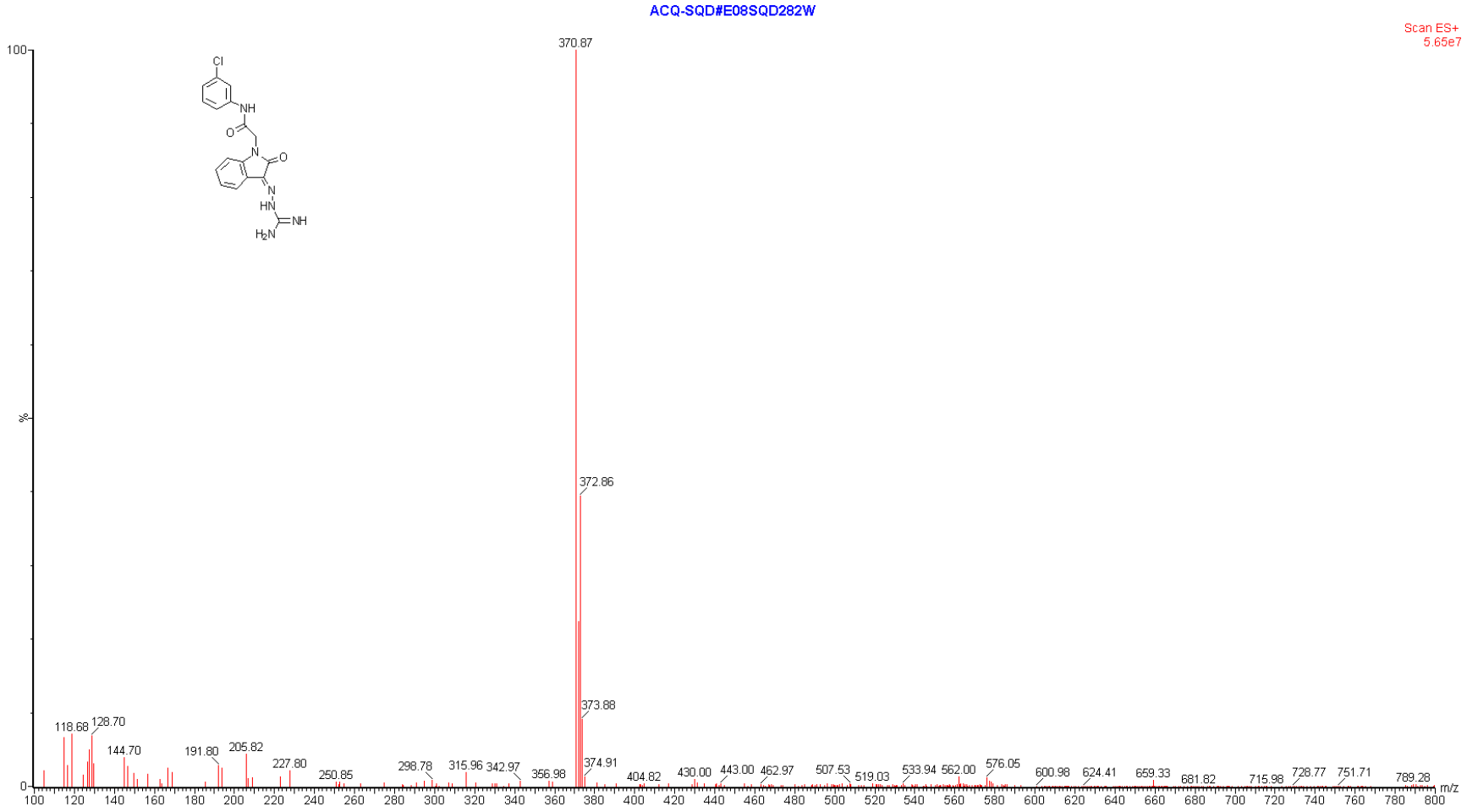
II-19 MS spectrum of compound GT10-18



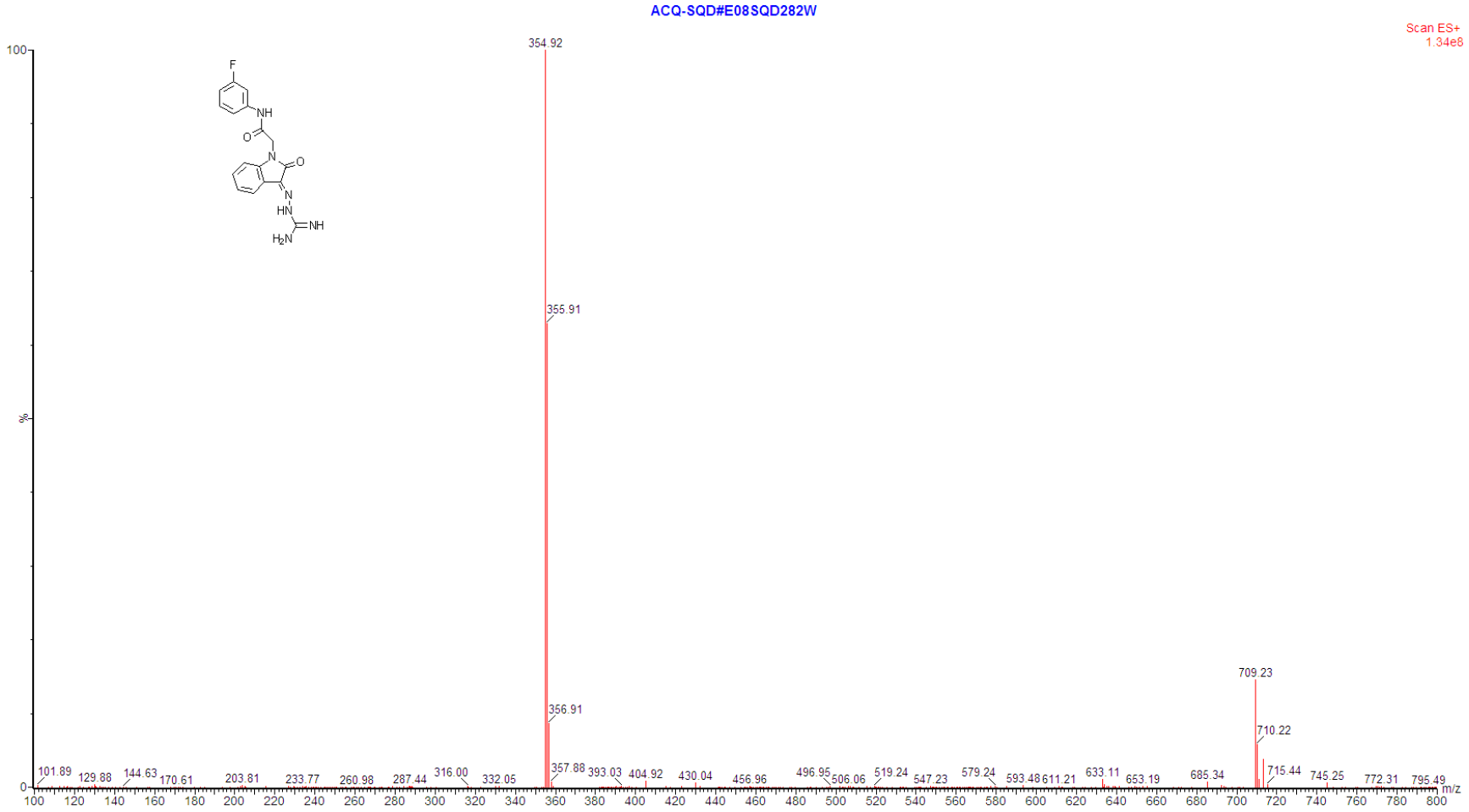
II-20 MS spectrum of compound GT10-19



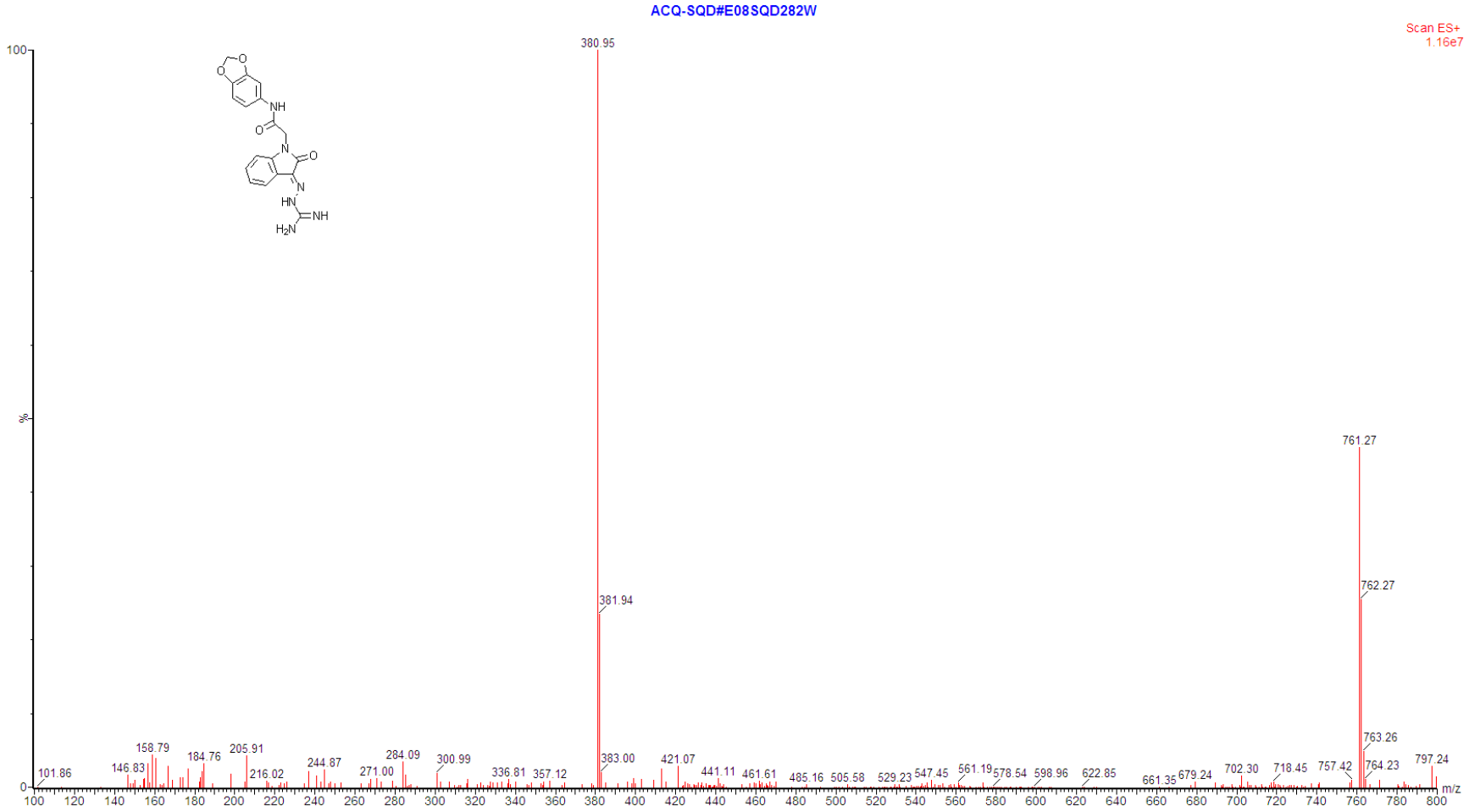
II-21 MS spectrum of compound GT10-20



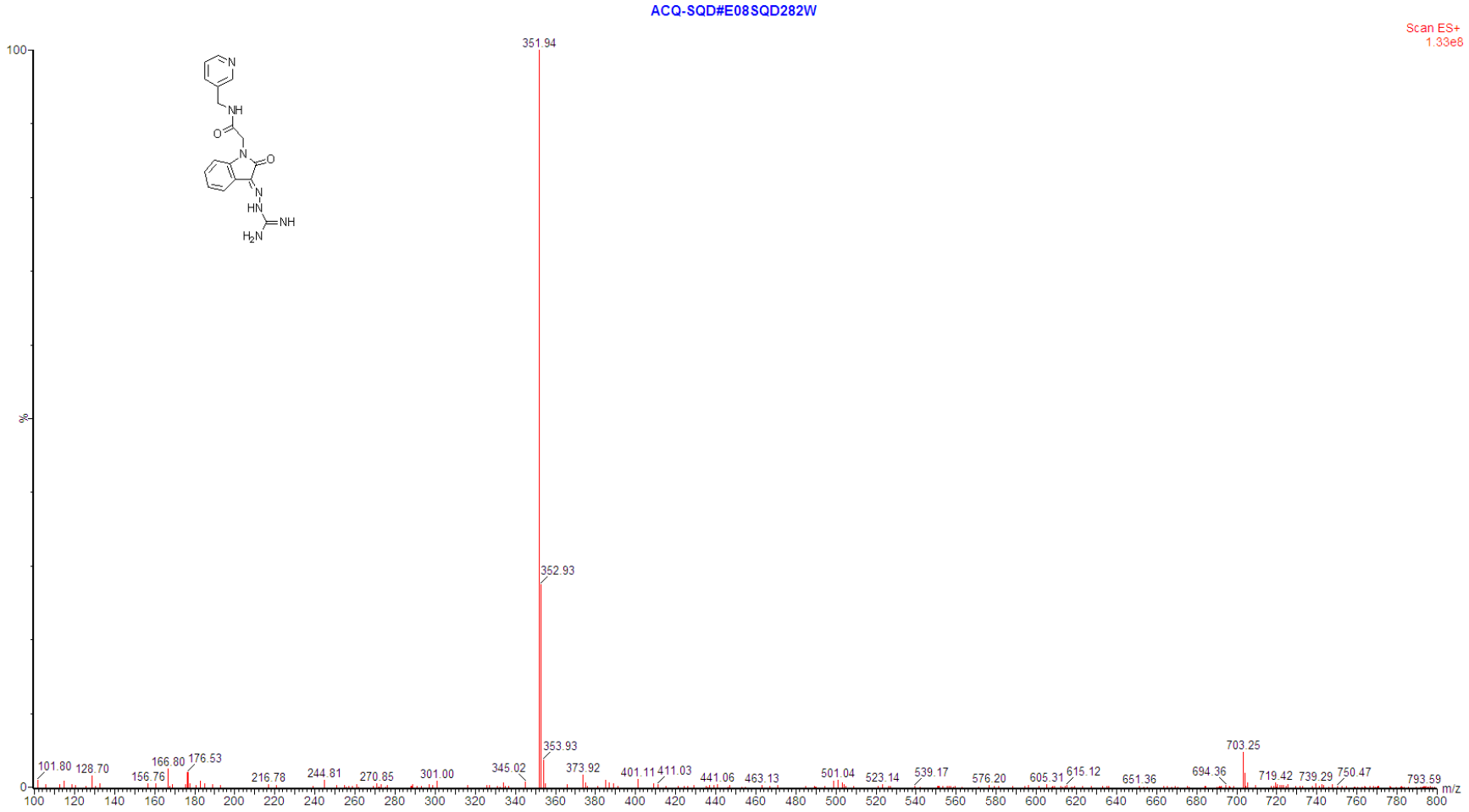
II-22 MS spectrum of compound GT10-21



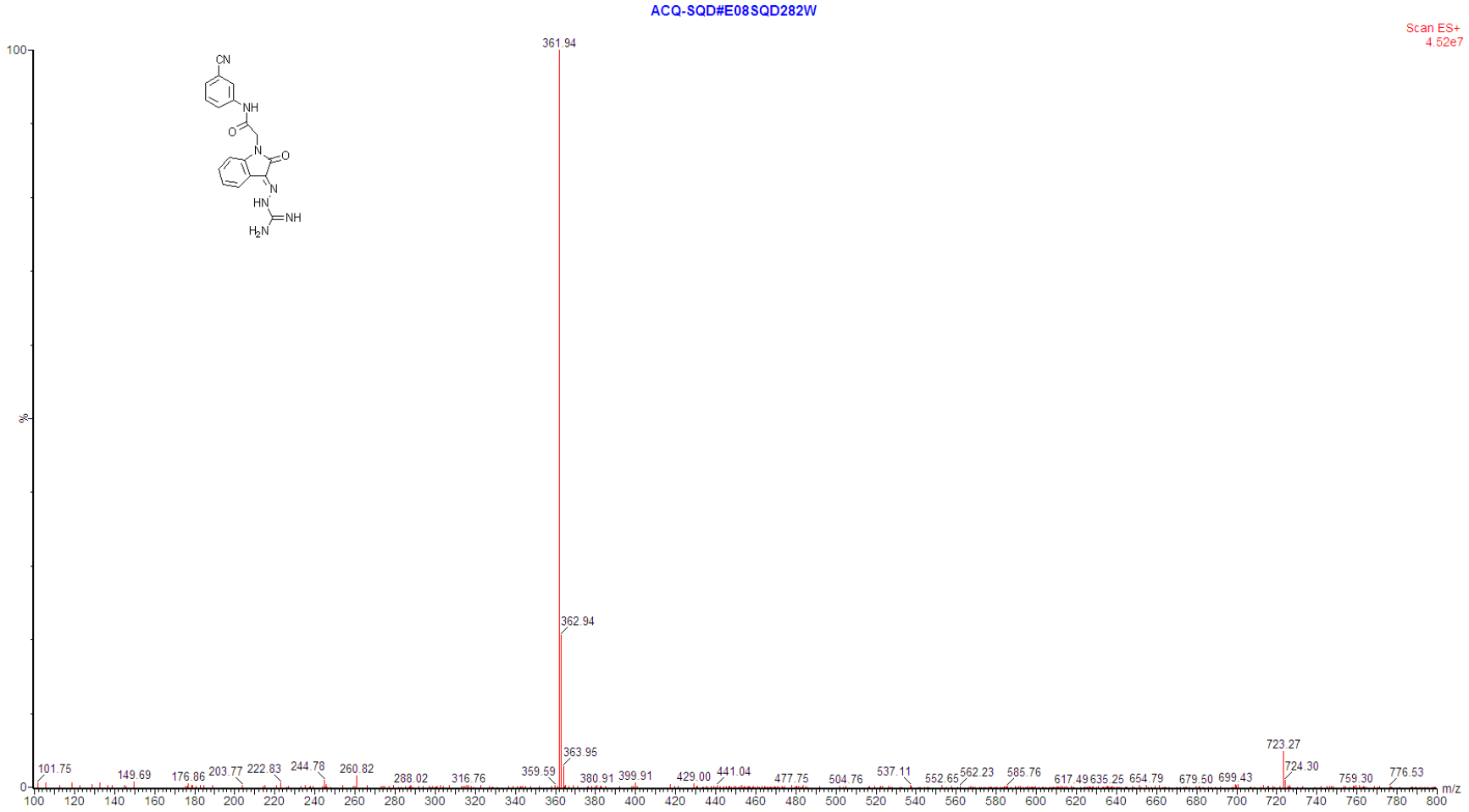
II-23 MS spectrum of compound GT10-22



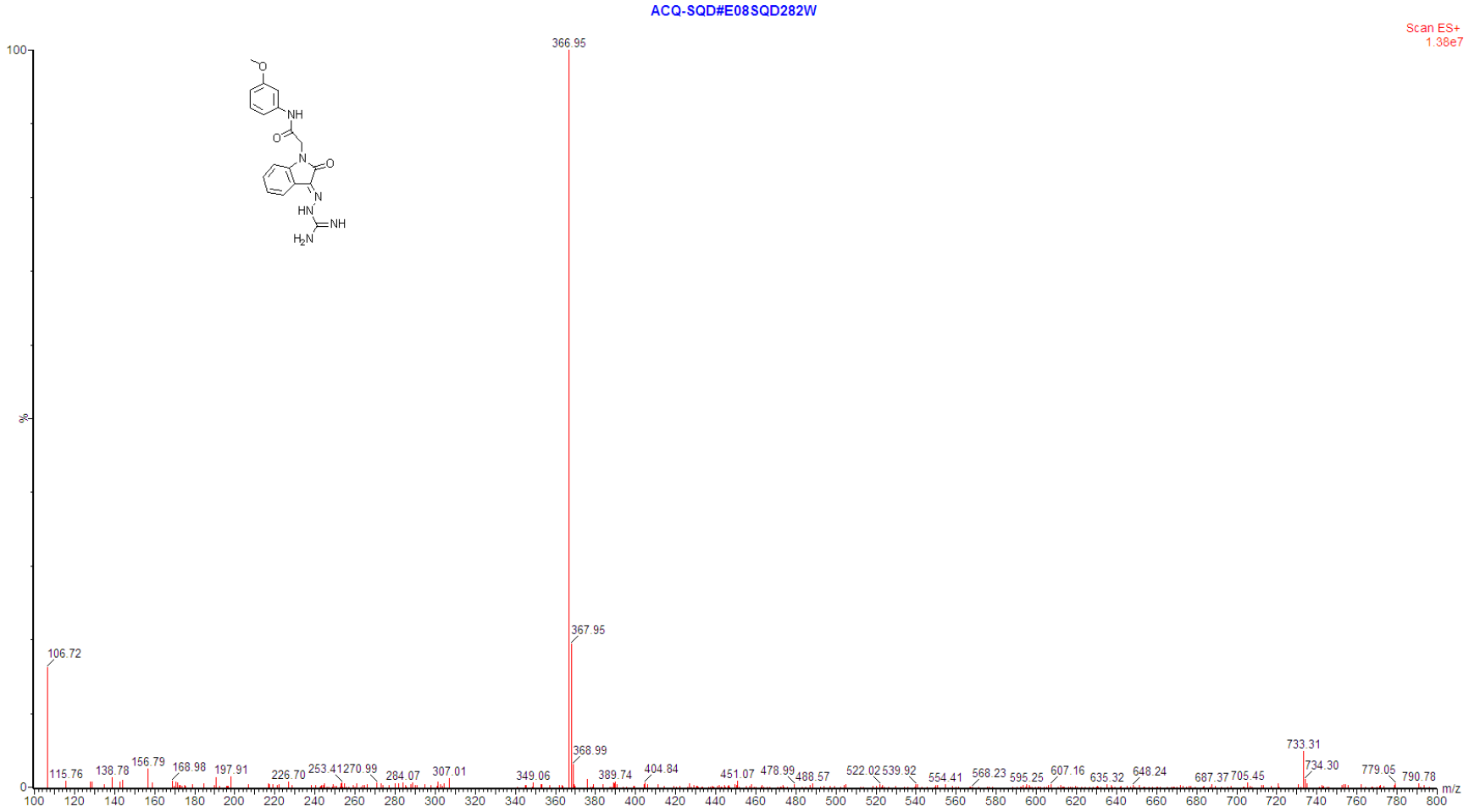
II-24 MS spectrum of compound GT10-23



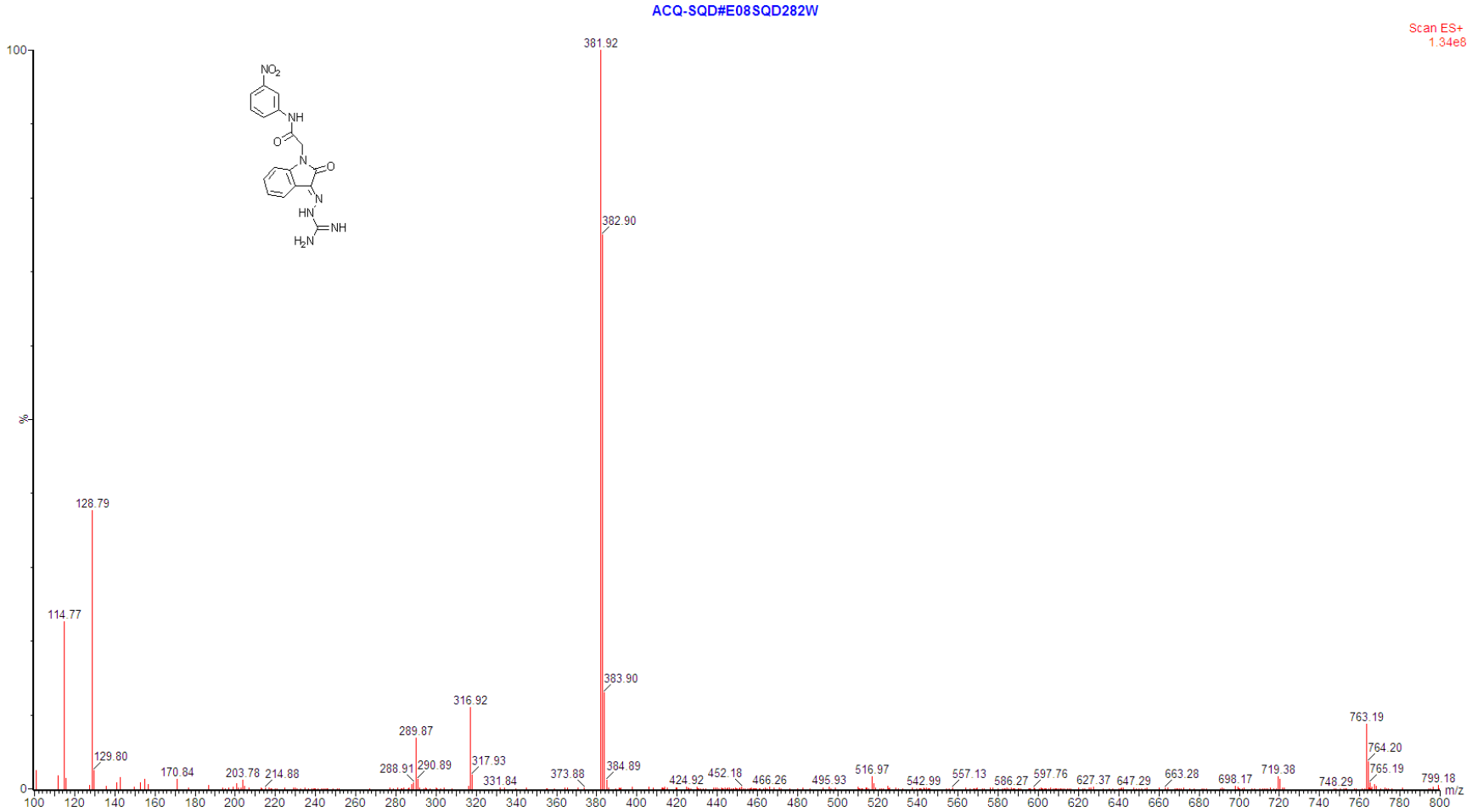
II-25 MS spectrum of compound GT10-24



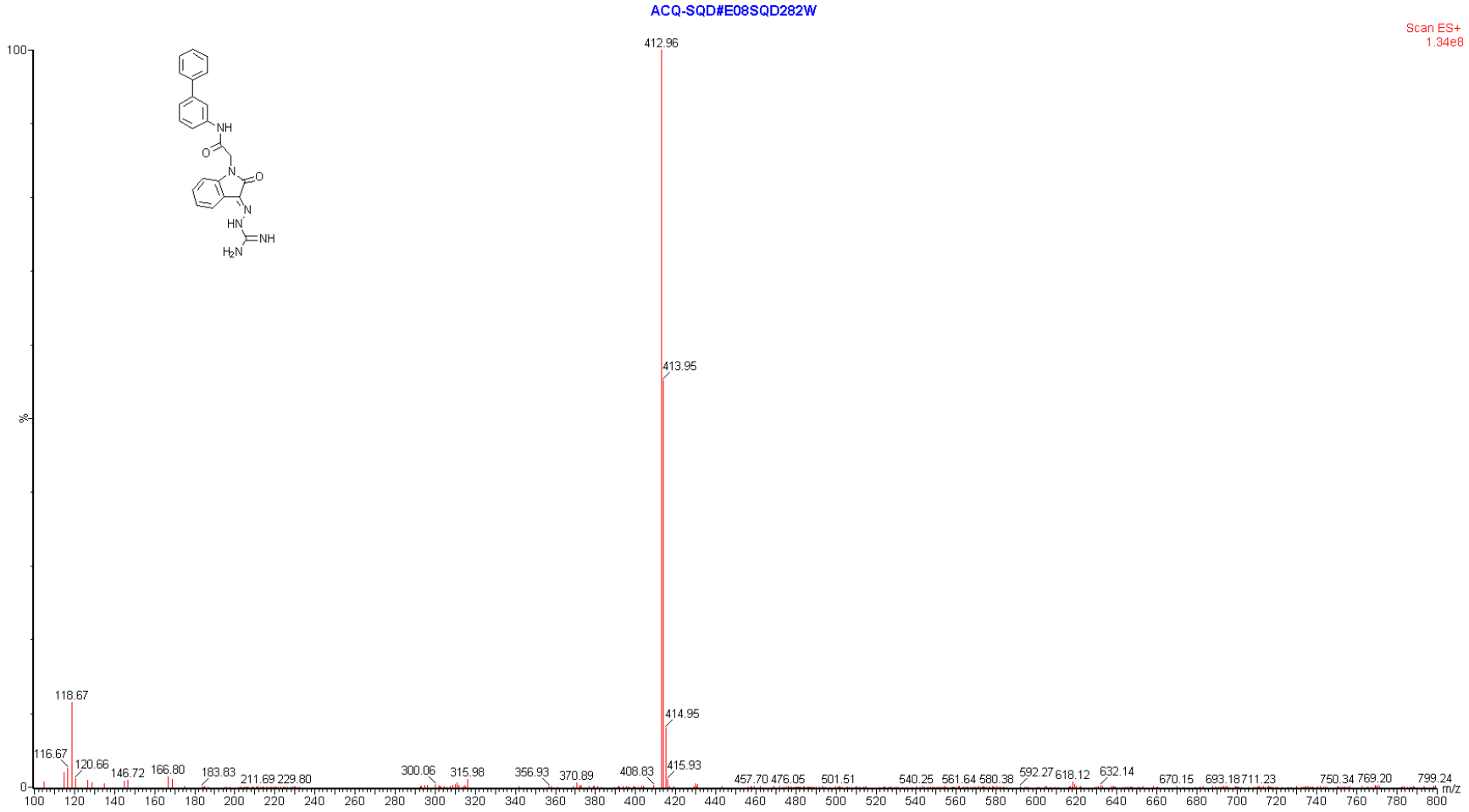
II-26 MS spectrum of compound GT10-25



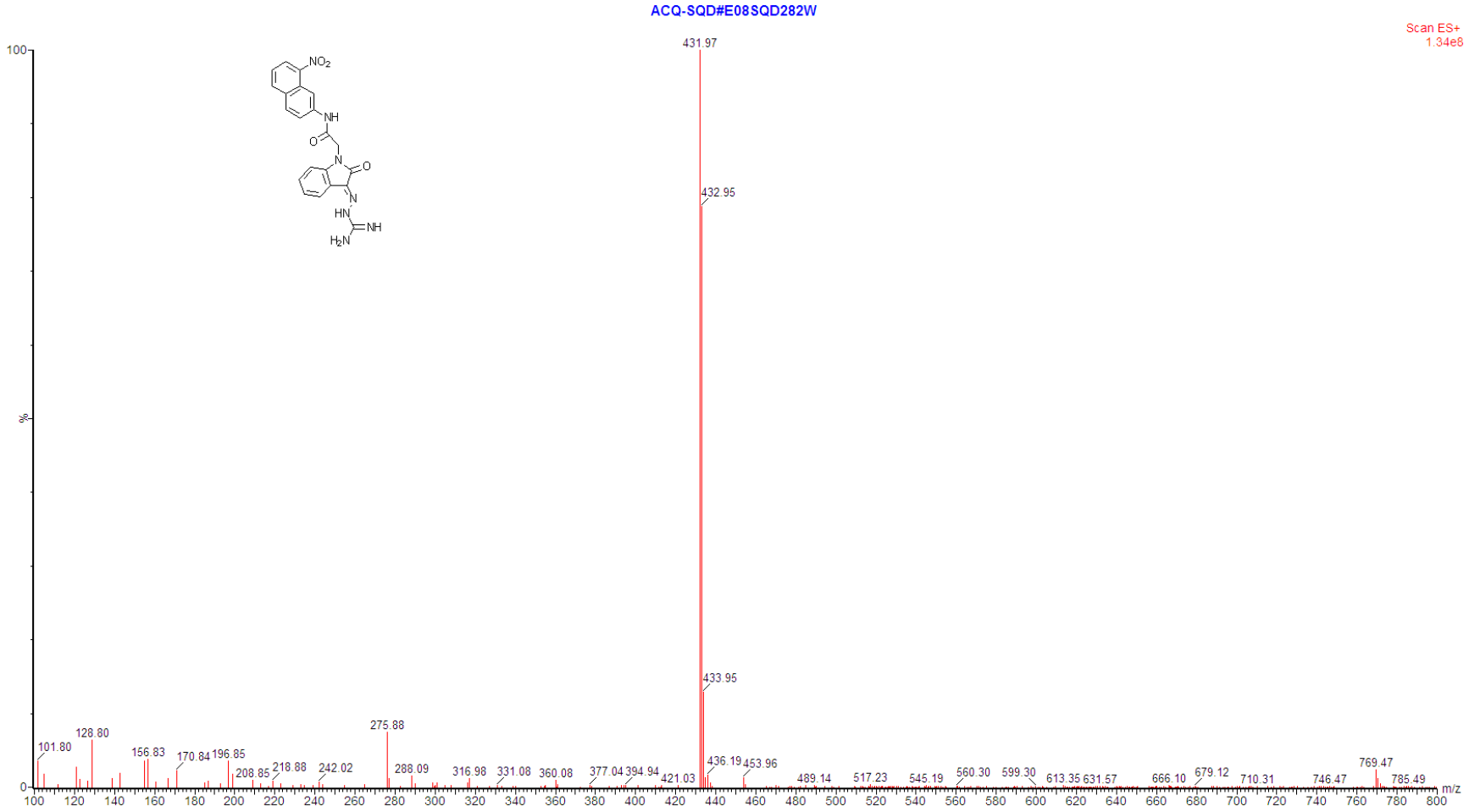
II-27 MS spectrum of compound GT10-26



II-28 MS spectrum of compound GT10-27



II-29 MS spectrum of compound GT10-28



II-30 MS spectrum of compound GT10-29

References

1. Mitsuoka T. Bifidobacteria and Their Role in Human Health. *Journal of Industrial Microbiology*, 1990, 6(4): 263-267.
2. Pereira D. I. A.; Gibson G. R. Cholesterol assimilation by lactic acid bacteria and bifidobacteria isolated from the human gut. *Applied and Environmental Microbiology*, 2002, 68(9): 4689-4693.
3. Murphy O. Non-polyol low-digestible carbohydrates: food applications and functional benefits. *British Journal of Nutrition*, 2001, 85: S47-S53.
4. Koul A.; Arnoult E.; Lounis N.; Guillemont J.; Andries K. The challenge of new drug discovery for tuberculosis. *Nature*, 2011, 469(7331): 483-490.
5. Chong C. P.; Street P. R. Pneumonia in the Elderly: A Review of Severity Assessment, Prognosis, Mortality, Prevention, and Treatment. *Southern Medical Journal*, 2008, 101(11): 1134-1140.
6. Walker C. L. F.; Rudan I.; Liu L.; Nair H.; Theodoratou E.; Bhutta Z. A.; O'Brien K. L.; Campbell H.; Black R. E. Childhood Pneumonia and Diarrhoea 1 Global burden of childhood pneumonia and diarrhoea. *Lancet*, 2013, 381(9875): 1405-1416.
7. Lewis K. Platforms for antibiotic discovery. *Nature Reviews Drug Discovery*, 2013, 12(5): 371-387.
8. Mehndiratta P. L.; Bhalla P. Typing of Methicillin resistant Staphylococcus aureus: A technical review. *Indian Journal of Medical Microbiology*, 2012, 30(1): 16-23.

9. Moellering R. C. MRSA: the first half century. *Journal of Antimicrobial Chemotherapy*, 2012, 67(1): 4-11.

10. Wright G. D. Antibiotic resistance in the environment: a link to the clinic? *Current Opinion in Microbiology*, 2010, 13(5): 589-594.

11. Barbosa T. M.; Levy S. B. The impact of antibiotic use on resistance development and persistence. *Drug Resistance Updates*, 2000, 3(5): 303-311.

12. Levy S. B.; Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. *Nature Medicine*, 2004, 10(12): S122-S129.

13. Bennett P. M. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *British Journal of Pharmacology*, 2008, 153: S347-S357.

14. Leclercq R.; Courvalin P. Resistance to glycopeptides in enterococci. *Clinical Infectious Diseases*, 1997, 24(4): 545-556.

15. Levy S. B. Active Efflux Mechanisms for Antimicrobial Resistance. *Antimicrobial Agents and Chemotherapy*, 1992, 36(4): 695-703.

16. Bradford P. A. Extended-spectrum beta-lactamases in the 21st century: Characterization, epidemiology, and detection of this important resistance threat. *Clinical Microbiology Reviews*, 2001, 14(4): 933-951.

17. Davies J.; Wright G. D. Bacterial resistance to aminoglycoside antibiotics. *Trends in Microbiology*, 1997, 5(6): 234-240.

18. Hawkey P. M. The origins and molecular basis of antibiotic resistance. *British Medical Journal*, 1998, 317(7159): 657-660.

19. Lazar K.; Walker S. Substrate analogues to study cell-wall biosynthesis and its inhibition. *Current Opinion in Chemical Biology*, 2002, 6(6): 786-793.

20. Baizman E. R.; Branstrom A. A.; Longley C. B.; Allanson N.; Sofia M. J.; Gange D.; Goldman R. C. Antibacterial activity of synthetic analogues based on the disaccharide structure of moenomycin, an inhibitor of bacterial transglycosylase. *Microbiology-Uk*, 2000, 146: 3129-3140.

21. van Heijenoort J. Recent advances in the formation of the bacterial peptidoglycan monomer unit. *Natural Product Reports*, 2001, 18(5): 503-519.

22. van Heijenoort J. Formation of the glycan chains in the synthesis of bacterial peptidoglycan. *Glycobiology*, 2001, 11(3): 25r-36r.

23. Lovering A. L.; Safadi S. S.; Strynadka N. C. J. Structural Perspective of Peptidoglycan Biosynthesis and Assembly. *Annual Review of Biochemistry*, 2012, 81: 451-478.

24. Reed P.; Veiga H.; Jorge A. M.; Terrak M.; Pinho M. G. Monofunctional Transglycosylases Are Not Essential for Staphylococcus aureus Cell Wall Synthesis. *Journal of Bacteriology*, 2011, 193(10): 2549-2556.

25. Basu J.; Chattopadhyay R.; Kundu M.; Chakrabarti P. Purification and Partial Characterization of a Penicillin-Binding Protein from Mycobacterium-Smegmatis. *Journal of Bacteriology*, 1992, 174(14): 4829-4832.

26. Popham D. L.; Young K. D. Role of penicillin-binding proteins in bacterial cell morphogenesis. *Current Opinion in Microbiology*, 2003, 6(6): 594-599.

27. Spratt B. G. Distinct penicillin binding proteins involved in the division, elongation, and shape of Escherichia coli K12. *Proceedings of the National Academy of Sciences of the United States of America*, 1975, 72(8): 2999-3003.

28. Davies J.; Davies D. Origins and Evolution of Antibiotic Resistance. *Microbiology and Molecular Biology Reviews*, 2010, 74(3): 417-433.

29. Roth S.; Mcguire E. J.; Roseman S. Evidence for Cell-Surface Glycosyltransferases - Their Potential Role in Cellular Recognition. *Journal of Cell Biology*, 1971, 51(2): 536-547.

30. Goffin C.; Ghuysen J. M. Multimodular penicillin binding proteins: An enigmatic family of orthologs and paralogs. *Microbiology and Molecular Biology Reviews*, 1998, 62(4): 1079-1093.

31. Huang C. Y.; Shih H. W.; Lin L. Y.; Tien Y. W.; Cheng T. J. R.; Cheng W. C.; Wong C. H.; Ma C. Crystal structure of Staphylococcus aureus transglycosylase in complex with a lipid II analog and elucidation of peptidoglycan synthesis mechanism.

Proceedings of the National Academy of Sciences of the United States of America, 2012, 109(17): 6496-6501.

32. Yuan Y. Q.; Fuse S.; Ostash B.; Sliz P.; Kahne D.; Walker S. Structural analysis of the contacts anchoring moenomycin to peptidoglycan glycosyltransferases and implications for antibiotic design. *ACS Chemical Biology*, 2008, 3(7): 429-436.

33. Lovering A. L.; de Castro L. H.; Lim D.; Strynadka N. C. J. Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis. *Science*, 2007, 315(5817): 1402-1405.

34. Sung M. T.; Lai Y. T.; Huang C. Y.; Chou L. Y.; Shih H. W.; Cheng W. C.; Wong C. H.; Ma C. Crystal structure of the membrane-bound bifunctional transglycosylase PBP1b from Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(22): 8824-8829.

35. Reynolds P. E. Structure, Biochemistry and Mechanism of Action of Glycopeptide Antibiotics. *European Journal of Clinical Microbiology and Infectious Diseases*, 1989, 8(11): 943-950.

36. Hammes W. P.; Neuhaus F. C. Mechanism of Action of Vancomycin - Inhibition of Peptidoglycan Synthesis in Gaffkya-Homari. *Antimicrobial Agents and Chemotherapy*, 1974, 6(6): 722-728.

37. Welzel P. Syntheses around the transglycosylation step in peptidoglycan biosynthesis. *Chemical Reviews*, 2005, 105(12): 4610-4660.

38. Cetinkaya Y.; Falk P.; Mayhall C. G. Vancomycin-resistant enterococci. *Clinical Microbiology Reviews*, 2000, 13(4): 686-707.

39. Courvalin P. Vancomycin resistance in gram-positive cocci. *Clinical Infectious Diseases*, 2006, 42: S25-S34.

40. Malabarba A.; Nicas T. I.; Thompson R. C. Structural modifications of glycopeptide antibiotics. *Medicinal Research Reviews*, 1997, 17(1): 69-137.

41. Goldman R. C.; Gange D. Inhibition of transglycosylation involved in bacterial peptidoglycan synthesis. *Current Medicinal Chemistry*, 2000, 7(8): 801-820.

42. Healy V. L.; Lessard I. A. D.; Roper D. I.; Knox J. R.; Walsh C. T. Vancomycin resistance in enterococci: reprogramming of the D-Ala-D-Ala ligases in bacterial peptidoglycan biosynthesis. *Chemistry & Biology*, 2000, 7(5): R109-R119.

43. Ritter T. K.; Wong C. H. Carbohydrate-based antibiotics: A new approach to tackling the problem of resistance. *Angewandte Chemie-International Edition*, 2001, 40(19): 3509-3533.

44. Jack R. W.; Jung G. Lantibiotics and microcins: polypeptides with unusual chemical diversity. *Current Opinion in Chemical Biology*, 2000, 4(3): 310-317.

45. Moll G. N.; Roberts G. C. K.; Konings W. N.; Driessen A. J. M. Mechanism of lantibiotic-induced pore-formation. *Antonie Van*

Leeuwenhoek International Journal of General and Molecular Microbiology, 1996, 69(2): 185-191.

46. Brotz H.; Sahl H. G. New insights into the mechanism of action of lantibiotics - diverse biological effects by binding to the same molecular target. *Journal of Antimicrobial Chemotherapy*, 2000, 46(1): 1-6.

47. Bierbaum G.; Sahl H. G. Lantibiotics: Mode of Action, Biosynthesis and Bioengineering. *Current Pharmaceutical Biotechnology*, 2009, 10(1): 2-18.

48. Brotz H.; Bierbaum G.; Leopold K.; Reynolds P. E.; Sahl H. G. The lantibiotic mersacidin inhibits peptidoglycan synthesis by targeting lipid II. *Antimicrobial Agents and Chemotherapy*, 1998, 42(1): 154-160.

49. He H. Y.; Shen B.; Korshalla J.; Siegel M. M.; Carter G. T. Isolation and structural elucidation of AC326-alpha, a new member of the moenomycin group. *Journal of Antibiotics*, 2000, 53(2): 191-195.

50. Pfaller M. A. Flavophospholipol use in animals: Positive implications for antimicrobial resistance based on its microbiologic properties. *Diagnostic Microbiology and Infectious Disease*, 2006, 56(2): 115-121.

51. Taylor J. G.; Li X. C.; Oberthur M.; Zhu W. J.; Kahne D. E. The total synthesis of moenomycin A. *Journal of the American Chemical Society*, 2006, 128(47): 15084-15085.

52. Welzel P.; Kunisch F.; Kruggel F.; Stein H.; Scherkenbeck J.; Hiltmann A.; Duddeck H.; Muller D.; Maggio J. E.; Fehlhaber H. W.; Seibert G.; Vanheijenoort Y.; Vanheijenoort J. Moenomycin-a - Minimum Structural Requirements for Biological-Activity. *Tetrahedron*, 1987, 43(3): 585-598.

53. Vogel S.; Stembera K.; Hennig L.; Findeisen M.; Giesa S.; Welzel P.; Lampilas M. Moenomycin analogues with modified lipid side chains from indium-mediated Barbier-type reactions. *Tetrahedron*, 2001, 57(19): 4139-4146.

54. Vogel S.; Stembera K.; Hennig L.; Findeisen M.; Giesa S.; Welzel P.; Tillier C.; Bonhomme C.; Lampilas M. Moenomycin analogues with long-chain amine lipid parts from reductive aminations. *Tetrahedron*, 2001, 57(19): 4147-4160.

55. El-Abadla N.; Lampilas M.; Hennig L.; Findeisen M.; Welzel P.; Muller D.; Markus A.; van Heijenoort J. Moenomycin A: The role of the methyl group in the moenuronamide unit and a general discussion of structure-activity relationships. *Tetrahedron*, 1999, 55(3): 699-722.

56. Ostash B.; Walker S. Moenomycin family antibiotics: chemical synthesis, biosynthesis, and biological activity. *Natural Product Reports*, 2010, 27(11): 1594-1617.

57. Ruhl T.; Daghish M.; Buchynskyy A.; Barche K.; Volke D.; Stembera K.; Kempin U.; Knoll D.; Hennig L.; Findeisen M.;

Oehme R.; Giesa S.; Ayala J.; Welzel P. Studies on the interaction of the antibiotic moenomycin A with the enzyme penicillin-binding protein 1b. *Bioorganic & Medicinal Chemistry*, 2003, 11(13): 2965-2981.

58. Anikin A.; Buchynskyy A.; Kempin U.; Stembera K.; Welzel P.; Lantzsch G. Membrane anchoring and intervesicle transfer of a derivative of the antibiotic moenomycin A. *Angewandte Chemie-International Edition*, 1999, 38(24): 3703-3707.

59. Gampe C. M.; Tsukamoto H.; Doud E. H.; Walker S.; Kahne D. Tuning the Moenomycin Pharmacophore To Enable Discovery of Bacterial Cell Wall Synthesis Inhibitors. *Journal of the American Chemical Society*, 2013, 135(10): 3776-3779.

60. Cheng T. J. R.; Sung M. T.; Liao H. Y.; Chang Y. F.; Chen C. W.; Huang C. Y.; Chou L. Y.; Wu Y. D.; Chen Y.; Cheng Y. S. E.; Wong C. H.; Ma C.; Cheng W. C. Domain requirement of moenomycin binding to bifunctional transglycosylases and development of high-throughput discovery of antibiotics. *Proceedings of the National Academy of Sciences of the United States of America*, 2008, 105(2): 431-436.

61. Kitchen D. B.; Decornez H.; Furr J. R.; Bajorath J. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nature Reviews Drug Discovery*, 2004, 3(11): 935-949.

62. Beddell C. R.; Goodford P. J.; Norrington F. E.; Wilkinson S.; Wootton R. Compounds Designed to Fit a Site of Known Structure in Human Hemoglobin. *British Journal of Pharmacology*, 1976, 57(2): 201-209.

63. Shoichet B. K.; McGovern S. L.; Wei B. Q.; Irwin J. J. Lead discovery using molecular docking. *Current Opinion in Chemical Biology*, 2002, 6(4): 439-446.

64. DiMasi J. A.; Hansen R. W.; Grabowski H. G. The price of innovation: new estimates of drug development costs. *Journal of Health Economics*, 2003, 22(2): 151-185.

65. Song C. M.; Lim S. J.; Tong J. C. Recent advances in computer-aided drug design. *Briefings in Bioinformatics*, 2009, 10(5): 579-591.

66. Oprea T. I.; Davis A. M.; Teague S. J.; Leeson P. D. Is there a difference between leads and drugs? A historical perspective. *Journal of Chemical Information and Computer Sciences*, 2001, 41(5): 1308-1315.

67. Taft C. A.; Da Silva V. B.; Da Silva C. H. T. D. Current topics in computer-aided drug design. *Journal of Pharmaceutical Sciences*, 2008, 97(3): 1089-1098.

68. Kapetanovic I. M. Computer-aided drug discovery and development (CADDD): In silico-chemico-biological approach. *Chemico-Biological Interactions*, 2008, 171(2): 165-176.

69. Cheng T. J.; Li Q. L.; Zhou Z. G.; Wang Y. L.; Bryant S. H. Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. *AAPS Journal*, 2012, 14(1): 133-141.

70. Collier G.; Vellore N. A.; Latour R. A.; Stuart S. J. Development of molecular simulation methods to accurately represent protein-surface interactions: Method assessment for the calculation of electrostatic effects. *Biointerphases*, 2009, 4(4): 57-64.

71. Cornell W. D.; Cieplak P.; Bayly C. I.; Gould I. R.; Merz K. M.; Ferguson D. M.; Spellmeyer D. C.; Fox T.; Caldwell J. W.; Kollman P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules *Journal of the American Chemical Society*, 1996, 118(9): 2309-2309.

72. Halgren T. A. Merck molecular force field .1. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 1996, 17(5-6): 490-519.

73. Brunger A. T.; Adams P. D. Molecular dynamics applied to X-ray structure refinement. *Accounts of Chemical Research*, 2002, 35(6): 404-412.

74. Alonso H.; Bliznyuk A. A.; Gready J. E. Combining docking and molecular dynamic simulations in drug design. *Medicinal Research Reviews*, 2006, 26(5): 531-568.

75. Jorgensen W. L.; TiradoRives J. Monte Carlo vs molecular dynamics for conformational sampling. *Journal of Physical Chemistry*, 1996, 100(34): 14508-14513.

76. Huang S. Y.; Grinter S. Z.; Zou X. Q. Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. *Physical Chemistry Chemical Physics*, 2010, 12(40): 12899-12908.

77. Huang S. Y.; Zou X. Q. Scoring and Lessons Learned with the CSAR Benchmark Using an Improved Iterative Knowledge-Based Scoring Function. *Journal of Chemical Information and Modeling*, 2011, 51(9): 2097-2106.

78. Bohm H. J. Prediction of binding constants of protein ligands: A fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *Journal of Computer-Aided Molecular Design*, 1998, 12(4): 309-323.

79. Wang R. X.; Lai L. H.; Wang S. M. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *Journal of Computer-Aided Molecular Design*, 2002, 16(1): 11-26.

80. Zheng Z.; Merz K. M. Development of the Knowledge-Based and Empirical Combined Scoring Algorithm (KECSA) To Score Protein-Ligand Interactions. *Journal of Chemical Information and Modeling*, 2013, 53(5): 1073-1083.

81. Muegge I.; Martin Y. C. A general and fast scoring function for protein-ligand interactions: A simplified potential approach. *Journal of Medicinal Chemistry*, 1999, 42(5): 791-804.

82. McInnes C. Virtual screening strategies in drug discovery. *Current Opinion in Chemical Biology*, 2007, 11(5): 494-502.

83. Lyne P. D. Structure-based virtual screening: an overview. *Drug Discovery Today*, 2002, 7(20): 1047-1055.

84. Celej M. S.; Montich C. G.; Fidelio G. D. Protein stability induced by ligand binding correlates with changes in protein flexibility. *Protein Science*, 2003, 12(7): 1496-1506.

85. Carlson H. A. Protein flexibility and drug design: how to hit a moving target. *Current Opinion in Chemical Biology*, 2002, 6(4): 447-452.

86. Knegtel R. M. A.; Kuntz I. D.; Oshiro C. M. Molecular docking to ensembles of protein structures. *Journal of Molecular Biology*, 1997, 266(2): 424-440.

87. Najmanovich R.; Kuttner J.; Sobolev V.; Edelman M. Side-chain flexibility in proteins upon ligand binding. *Proteins-Structure Function and Genetics*, 2000, 39(3): 261-268.

88. Mangoni R.; Roccatano D.; Di Nola A. Docking of flexible ligands to flexible receptors in solution by molecular dynamics simulation. *Proteins-Structure Function and Bioinformatics*, 1999, 35(2): 153-162.

89. Schames J. R.; Henchman R. H.; Siegel J. S.; Sotriffer C. A.; Ni H. H.; McCammon J. A. Discovery of a novel binding trench in HIV integrase. *Journal of Medicinal Chemistry*, 2004, 47(8): 1879-1881.

90. Zhou Y.; Peng H.; Ji Q.; Qi J.; Zhu Z. P.; Yang C. Z. Discovery of small molecule inhibitors of integrin alpha v beta 3 through structure-based virtual screening. *Bioorganic & Medicinal Chemistry Letters*, 2006, 16(22): 5878-5882.

91. Bonacci T. M.; Mathews J. L.; Yuan C. J.; Lehmann D. M.; Malik S.; Wu D. Q.; Font J. L.; Bidlack J. M.; Smrcka A. V. Differential targeting of G beta gamma-subunit signaling with small molecules. *Science*, 2006, 312(5772): 443-446.

92. Schapira M.; Raaka B. M.; Samuels H. H.; Abagyan R. Rational discovery of novel nuclear hormone receptor antagonists. *Proceedings of the National Academy of Sciences of the United States of America*, 2000, 97(3): 1008-1013.

93. Siddiquee K.; Zhang S.; Guida W. C.; Blaskovich M. A.; Greedy B.; Lawrence H. R.; Yip M. L. R.; Jove R.; McLaughlin M. M.; Lawrence N. J.; Sebti S. M.; Turkson J. Selective chemical probe inhibitor of Stat3, identified through structure-based virtual screening, induces antitumor activity. *Proceedings of the National Academy of Sciences of the United States of America*, 2007, 104(18): 7391-7396.

94. Perola E.; Xu K.; Kollmeyer T. M.; Kaufmann S. H.; Prendergast F. G.; Pang Y. P. Successful virtual screening of a chemical database for farnesyltransferase inhibitor leads. *Journal of Medicinal Chemistry*, 2000, 43(3): 401-408.

95. Martin Y. C.; Kofron J. L.; Traphagen L. M. Do structurally similar molecules have similar biological activity? *Journal of Medicinal Chemistry*, 2002, 45(19): 4350-4358.

96. Hamza A.; Wei N. N.; Zhan C. G. Ligand-Based Virtual Screening Approach Using a New Scoring Function. *Journal of Chemical Information and Modeling*, 2012, 52(4): 963-974.

97. von Korff M.; Freyss J.; Sander T. Flexophore, a new versatile 3D pharmacophore descriptor that considers molecular flexibility. *Journal of Chemical Information and Modeling*, 2008, 48(4): 797-810.

98. Ahlstrom M. M.; Ridderstrom M.; Luthman K.; Zamora I. Virtual screening and scaffold hopping based on GRID molecular interaction fields. *Journal of Chemical Information and Modeling*, 2005, 45(5): 1313-1323.

99. Krishnan V. V. Ligand screening by saturation-transfer difference (STD) NMR spectroscopy. *Current Analytical Chemistry*, 2005, 1(3): 307-320.

100. Viegas A.; Manso J.; Nobrega F. L.; Cabrita E. J. Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for

Ligand Screening and Characterization of Protein Binding. *Journal of Chemical Education*, 2011, 88(7): 990-994.

101. Wagstaff J. L.; Taylor S. L.; Howard M. J. Recent developments and applications of saturation transfer difference nuclear magnetic resonance (STD NMR) spectroscopy. *Molecular BioSystems*, 2013, 9(4): 571-577.

102. Williamson M. (2006) The Nuclear Overhauser Effect. *Modern Magnetic Resonance*, ed Webb G (Springer Netherlands), pp 409-412.

103. Wang Y. S.; Liu D. J.; Wyss D. F. Competition STD NMR for the detection of high-affinity ligands and NMR-based screening. *Magnetic Resonance in Chemistry*, 2004, 42(6): 485-489.

104. Angulo J.; Enriquez-Navas P. M.; Nieto P. M. Ligand-Receptor Binding Affinities from Saturation Transfer Difference (STD) NMR Spectroscopy: The Binding Isotherm of STD Initial Growth Rates. *Chemistry-a European Journal*, 2010, 16(26): 7803-7812.

105. Hiraishi N.; Tochio N.; Kigawa T.; Otsuki M.; Tagami J. Monomer-Collagen Interactions Studied by Saturation Transfer Difference NMR. *Journal of Dental Research*, 2013, 92(3): 284-288.

106. Chandrakala B.; Shandil R. K.; Mehraj U.; Ravishankar S.; Kaur P.; Usha V.; Joe B.; deSousa S. M. High-throughput screen for inhibitors of transglycosylase and/or transpeptidase activities of

Escherichia coli penicillin binding protein 1b. *Antimicrobial Agents and Chemotherapy*, 2004, 48(1): 30-40.

107. Donovan R. S.; Datti A.; Baek M. G.; Wu Q. Q.; Sas I. J.; Korczak B.; Berger E. G.; Roy R.; Dennis J. W. A solid-phase glycosyltransferase assay for high-throughput screening in drug discovery research. *Glycoconjugate Journal*, 1999, 16(10): 607-615.

108. Hu Y.; Heim J. S.; Chen L.; Ginsberg C.; Gross B.; Kraybill B.; Tiyanont K.; Fang X.; Wu T.; Walker S. Identification of selective inhibitors for the glycosyltransferase via high-throughput murG screening. *Chemistry & Biology*, 2004, 11(5): 703-711.

109. Doman T. N.; McGovern S. L.; Witherbee B. J.; Kasten T. P.; Kurumbail R.; Stallings W. C.; Connolly D. T.; Shoichet B. K. Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *Journal of Medicinal Chemistry*, 2002, 45(11): 2213-2221.

110. Abagyan R.; Totrov M.; Kuznetsov D. Icm - a New Method for Protein Modeling and Design - Applications to Docking and Structure Prediction from the Distorted Native Conformation. *Journal of Computational Chemistry*, 1994, 15(5): 488-506.

111. Argos P.; Abagyan R. The Protein-Folding Problem - Finding a Few Minimums in a near Infinite Space. *Computers & Chemistry*, 1994, 18(3): 225-231.

112. Abagyan R.; Totrov M. Biased Probability Monte-Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins. *Journal of Molecular Biology*, 1994, 235(3): 983-1002.

113. Schapira M.; Abagyan R.; Totrov M. Nuclear hormone receptor targeted virtual screening. *Journal of Medicinal Chemistry*, 2003, 46(14): 3045-3059.

114. Lovering A. L.; De Castro L.; Strynadka N. C. J. Identification of Dynamic Structural Motifs Involved in Peptidoglycan Glycosyltransfer. *Journal of Molecular Biology*, 2008, 383(1): 167-177.

115. Terrak M.; Sauvage E.; Derouaux A.; Dehareng D.; Bouhss A.; Breukink E.; Jeanjean S.; Nguyen-Disteche M. Importance of the conserved residues in the peptidoglycan glycosyltransferase module of the class a penicillin-binding protein 1b of Escherichia coli. *Journal of Biological Chemistry*, 2008, 283(42): 28464-28470.

116. Breukink E.; de Kruijff B. Lipid II as a target for antibiotics. *Nature Reviews Drug Discovery*, 2006, 5(4): 321-332.

117. Neves M. A. C.; Totrov M.; Abagyan R. Docking and scoring with ICM: the benchmarking results and strategies for improvement. *Journal of Computer-Aided Molecular Design*, 2012, 26(6): 675-686.

118. Lipinski C. A.; Lombardo F.; Dominy B. W.; Feeney P. J. Experimental and computational approaches to estimate solubility

and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 2001, 46(1-3): 3-26.

119. Totrov M.; Abagyan R. Flexible protein-ligand docking by global energy optimization in internal coordinates. *Proteins-Structure Function and Genetics*, 1997: 215-220.

120. An J. H.; Totrov M.; Abagyan R. Pocketome via comprehensive identification and classification of ligand binding envelopes. *Molecular and Cellular Proteomics*, 2005, 4(6): 752-761.

121. Bradford M. M. Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Analytical Biochemistry*, 1976, 72(1-2): 248-254.

122. da Silva J. F. M.; Garden S. J.; Pinto A. C. The chemistry of isatins: a review from 1975 to 1999. *Journal of the Brazilian Chemical Society*, 2001, 12(3): 273-324.

123. Medvedev A.; Igosheva N.; Crumeyrolle-Arias M.; Glover V. Isatin: Role in stress and anxiety. *Stress-the International Journal on the Biology of Stress*, 2005, 8(3): 175-183.

124. Sridhar S. K.; Saravanan M.; Ramesh A. Synthesis and antibacterial screening of hydrazones, Schiff and Mannich bases of isatin derivatives. *European Journal of Medicinal Chemistry*, 2001, 36(7-8): 615-625.

125. Pandeya S. N.; Sriram D.; Nath G.; De Clercq E. Synthesis and antimicrobial activity of Schiff and Mannich bases of isatin and its derivatives with pyrimidine. *Farmaco,* 1999, 54(9): 624-628.

126. Pirrung M. C.; Pansare S. V.; Das Sarma K.; Keith K. A.; Kern E. R. Combinatorial optimization of isatin-beta-thiosemicarbazones as anti-poxvirus agents. *Journal of Medicinal Chemistry*, 2005, 48(8): 3045-3050.

127. Arafa R. K.; Hegazy G. H.; Piazza G. A.; Abadi A. H. Synthesis and in vitro antiproliferative effect of novel quinoline-based potential anticancer agents. *European Journal of Medicinal Chemistry,* 2013, 63: 826-832.

128. Morrison D. C.; Lee H. P. C. A New Preparation of 8-Nitro-2-naphthylamine. *The Journal of Organic Chemistry,* 1962, 27(9): 3336-3337.

129. Mok N. Y.; Chadwick J.; Kellett K. A. B.; Hooper N. M.; Johnson A. P.; Fishwick C. W. G. Discovery of novel non-peptide inhibitors of BACE-1 using virtual high-throughput screening. *Bioorganic & Medicinal Chemistry Letters*, 2009, 19(23): 6770-6774.

130. Modi N. R.; Shah R. J.; Patel M. J.; Suthar M.; Chauhan B. F.; Patel L. J. Design, synthesis, and QSAR study of novel 2-(2,3-dioxo-2,3-dihydro-1H-indol-1-yl)-N-phenylacetamide derivatives as cytotoxic agents. *Medicinal Chemistry Research*, 2011, 20(5): 615-625.

131. Tian X.; Wang L. Y.; Xia S. A.; Li Z. B.; Liu X. H.; Yuan Y.; Fang L.; Zuo H. Synthesis of 2H-benzo[b][1,4]oxazin-3(4H)-one derivatives as platelet aggregation inhibitors. *Bioorganic & Medicinal Chemistry Letters*, 2012, 22(1): 204-206.

132. Ju K. S.; Parales R. E. Nitroaromatic Compounds, from Synthesis to Biodegradation. *Microbiology and Molecular Biology Reviews*, 2010, 74(2): 250-272.

133. Candido-Bacani P. D.; Mori M. P.; Calvo T. R.; Vilegas W.; Varanda E. A.; Colus I. M. D. In Vitro Assessment of the Cytotoxic, Apoptotic, and Mutagenic Potentials of Isatin. *Journal of Toxicology and Environmental Health, Part A: Current Issues*, 2013, 76(6): 354-362.

134. Malpani Y.; Achary R.; Kim S. Y.; Jeong H. C.; Kim P.; Han S. B.; Kim M.; Lee C. K.; Kim J. N.; Jung Y. S. Efficient synthesis of 3H,3'H-spiro[benzofuran-2,1'-isobenzofuran]-3,3'-dione as novel skeletons specifically for influenza virus type B inhibition. *European Journal of Medicinal Chemistry*, 2013, 62: 534-544.

135. Lipinski C. A. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies*, 2004, 1(4): 337-341.