



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS
QUÍMICAS**

RELACIONES ESTRUCTURA–ACTIVIDAD, DISEÑO, SÍNTESIS Y
ACTIVIDAD ANTIPROTOZOARIA DE DERIVADOS DE BENCIMIDAZOL

TESIS
PARA OPTAR POR EL GRADO DE

DOCTOR EN CIENCIAS

PRESENTA

M. en C. JAIME PÉREZ VILLANUEVA



TUTOR: Dr. Rafael Castillo Bocanegra
Facultad de Química

Septiembre 2013



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

Mi más sincero agradecimiento al Dr. Rafael Castillo Bocanegra por sus enseñanzas y consejos tanto en el área de la química farmacéutica como personales. Además, me es grato reconocer que gracias a él ha crecido enormemente mi interés por la investigación en esta disciplina.

A la M. en C. María Alicia Hernández Campos, al Dr. Francisco Hernández Luis, por sus valiosas aportaciones en el desarrollo de este trabajo.

Al Dr. José Luis Medina Franco y a la Dra. Karina Martínez Mayorga les agradezco permitirme colaborar con su grupo. Estoy muy agradecido por todas las enseñanzas, consejos y apoyos que me brindaron durante mi estancia en *Torrey Pines Institute for Molecular Studies*, en Florida, EUA.

Agradezco a los Doctores Fabián Harvey López Vallejo, Tomas R. Caulfield, Kendall Byler, Austin Yongye y a los Maestros Antonio Romo Mancillas, Carlos Méndez Cuesta y Oscar Méndez Lucio por todo su apoyo, consejos y enseñanzas.

Agradezco a la Dra. Lilián Yépez Mulia y a la Bióloga Amparo Tapia, del Centro Médico Nacional Siglo XXI del IMSS, por la determinación de la actividad antiparasitaria de los compuestos preparados. Su participación ha sido crucial para enriquecer este trabajo.

Al Dr. Enrique Ramón Ángeles Anguiano por sus consejos y sugerencias durante el desarrollo de este trabajo.

Al personal de la USAI de la Facultad de Química, UNAM, por la determinación de los espectros de los compuestos sintetizados en este trabajo.

A todos mis compañeros del laboratorio 122 por sus consejos, aportaciones y amistad.

Gracias a los miembros del jurado por sus valiosas correcciones y observaciones en esta tesis.

Al Consejo Nacional de Ciencia y Tecnología por el apoyo como becario durante mis estudios de doctorado (número de becario 173896) y a los proyectos DGAPA IN 211806 y CONACyT 80093 por el financiamiento de este trabajo.

JURADO ASIGNADO

Presidente	Dr. Francisco Yuste López
Vocal	Dra. Lilián Yépez Mulia
Vocal	Dr. Rubén Trinidad Sánchez Obregón
Vocal	Dr. Francisco Hernández Luis
Secretario	Dra. Karina Martínez Mayorga

La mayor parte del presente trabajo se realizó en la Facultad de Química, Departamento de Farmacia, laboratorio L-122 del conjunto E de la Universidad Nacional Autónoma de México, bajo la dirección del Dr. Rafael Castillo Bocanegra. Adicionalmente, una parte importante, que corresponde a los estudios de quimioinformática, se realizó en *Torrey Pines Institute for Molecular Studies*, en Florida, EUA, bajo la dirección del Dr. José Luis Medina Franco. Las evaluaciones antiparasitarias se realizaron en la Unidad de Investigación Médica en Enfermedades Infecciosas y Parasitarias del Centro Médico Nacional, IMSS, bajo la dirección de la Dra. Lilián Yépez Mulia.

Los resultados de esta tesis dieron lugar a las siguientes publicaciones:

- Jaime Pérez-Villanueva, Radleigh Santos, Alicia Hernández-Campos, Marc A. Giulianotti, Rafael Castillo, José L. Medina-Franco. Towards a systematic characterization of the activity landscape of benzimidazole derivatives. *Bioorganic & Medicinal Chemistry*, 18 (2010) 7380–7391.
- Jaime Pérez-Villanueva, Radleigh Santos, Alicia Hernández-Campos, Marc A. Giulianotti, Rafael Castillo, José L. Medina-Franco. Structure-activity relationships of benzimidazole derivatives as antiparasitic agents: Dual activity-difference (DAD) maps. *MedChemComm*, 2 (2011) 44–49.
- Jaime Pérez-Villanueva, José L. Medina-Franco, Thomas R. Caulfield, Alicia Hernández-Campos, Francisco Hernández-Luis, Lilián Yépez-Mulia, Rafael Castillo. Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) of some benzimidazole derivatives with trichomonocidal activity. *European Journal of Medicinal Chemistry*, 46 (2011) 3499–3508.
- Jaime Pérez-Villanueva, Antonio Romo-Mancillas, Alicia Hernández-Campos, Lilián Yépez-Mulia, Francisco Hernández-Luis, Rafael Castillo. Antiprotozoal activity of proton-pump inhibitors. *Bioorganic & Medicinal Chemistry Letters*, 21 (2011) 7351–7354.
- Jaime Pérez-Villanueva, Alicia Hernández-Campos, Lilián Yépez-Mulia, Carlos Méndez-Cuesta, Oscar Méndez-Lucio, Francisco Hernández-Luis, Rafael Castillo. Synthesis and antiprotozoal activity of novel 2-[[2-(1*H*-imidazol-1-yl)ethyl]sulfonyl]-1*H*-benzimidazole derivatives. *Bioorganic & Medicinal Chemistry Letters*, 23 (2013) 4221–4224.

Los resultados de esta tesis fueron presentados en los siguientes congresos:

- 5ª Reunión de la Academia Mexicana de Química Orgánica, 28–29 de mayo de 2009, Zacatecas, Zac.
Análisis comparativo de campos moleculares y análisis comparativo de índices de similitud de derivados del bencimidazol actividad tricomonocida
- 44º Congreso Mexicano de Química, 26–30 de septiembre de 2009, Puebla, Pue.
Análisis CoMFA y CoMSIA de derivados del bencimidazol actividad tricomonocida
- 239th National Meeting, 21–25 de marzo de 2010, San Francisco, CA.
Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) of some benzimidazole derivatives with trichomonocidal activity
- QuimiUNAM 2010, 24–26 de noviembre de 2010, Ciudad Universitaria, D.F.
Relaciones estructura-actividad de derivados del bencimidazol con actividad antiprotozoaria: mapas de diferencia de actividad dual
- 7ª Reunión de la Academia Mexicana de Química Orgánica, 4–8 de abril de 2011, Cuernavaca, Mor.
Síntesis y actividad antiprotozoaria de nuevos derivados del bencimidazol
- 46º Congreso Mexicano de Química, 10–14 de septiembre de 2011, Querétaro, Qro.
Síntesis y actividad antiprotozoaria de nuevos derivados del bencimidazol
- 241th National Meeting, 27–31 de marzo de 2011, Anaheim, CA.
Antiprotozoal activity landscape of benzimidazole derivatives: dual activity-similarity maps
- 244th National Meeting, 19–23 de agosto de 2012, Philadelphia, PA.
Synthesis and antiprotozoal activity of some 2- $\{[2-(1H\text{-imidazol-1-yl})\text{ethyl}]\text{sulfanyl}\}$ - $1H\text{-benzimidazole}$ derivatives

ÍNDICE GENERAL

	Pág.
Resumen	i
Abstract	iii
1. Introducción	1
2. Antecedentes	2
2.1. <i>Enfermedades parasitarias</i>	2
2.2. <i>Bencimidazoles como agentes antiprotozoarios</i>	3
2.3. <i>Diseño de fármacos asistido por computadora y su aplicación en la búsqueda de agentes antiprotozoarios</i>	8
2.3.1. <i>Diseño de fármacos asistido por computadora</i>	8
2.3.2. <i>Estudios QSAR-3D previos sobre bencimidazoles con actividad antiprotozoaria</i>	10
2.3.3. <i>Panorama de actividad y su utilidad en la construcción de modelos QSAR</i>	12
3. Justificación	15
4. Hipótesis	15
5. Objetivos	15
5.1. <i>Objetivo general</i>	15
6. Metodología	17
6.1. <i>Panorama de actividad tricomonocida y giardicida de derivados de bencimidazol</i>	17
6.1.1. <i>Base de datos</i>	17
6.1.2. <i>Búsqueda conformacional</i>	17
6.1.3. <i>Similitud estructural 2D y 3D</i>	18
6.1.4. <i>Similitud usando propiedades calculadas</i>	18
6.1.5. <i>Similitud en actividad</i>	19
6.1.6. <i>Mapas SAS</i>	20
6.1.7. <i>Mapas SAS consenso</i>	20
6.1.8. <i>Mapas DAD</i>	21
6.2. <i>Relaciones cuantitativas estructura-actividad de derivados de bencimidazol con actividad tricomonocida</i>	21
6.2.1. <i>Base de datos y actividad biológica</i>	21
6.2.2. <i>Selección del grupo de prueba y de trabajo</i>	21
6.2.3. <i>Preparación de estructuras</i>	22
6.2.4. <i>Alineamiento molecular</i>	22
6.2.5. <i>Selección de formas tautoméricas</i>	23
6.2.6. <i>Metodología CoMFA</i>	23
6.2.7. <i>Metodología CoMSIA</i>	24
6.2.8. <i>Análisis de mínimos cuadrados parciales</i>	24
6.2.9. <i>Validación</i>	25

6.3. Actividad antiprotozoaria de los inhibidores de la bomba de protones (PPIs)	26
6.3.1. Modelado molecular	26
6.3.2. Obtención de los compuestos	26
6.3.3. Evaluaciones biológicas	26
6.4. Diseño, síntesis y actividad antiprotozoaria de nuevos derivados de bencimidazol	27
6.4.1. Moléculas diseñadas	27
6.4.2. Síntesis de los compuestos diseñados	28
6.4.2.1. 1-(2-Cloroetil)-1H-imidazol	28
6.4.2.2. 1,2-Diaminobencenos 1–5 (Serie 1)	29
6.4.2.3. 1,2-Diaminobencenos 6–11 (Serie 2)	29
6.4.2.3. 1,2-Diaminobencenos 12–17 (Serie 3)	31
7. Resultados y discusión	32
7.1. Panorama de actividad tricomonocida y giardicida de derivados de bencimidazol	32
7.1.1. Estudio basado en 32 compuestos	35
7.1.1.1. Representaciones de similitud estructural	35
7.1.1.2. Distribución de las representaciones de similitud estructural	35
7.1.1.3. Correlación entre similitudes moleculares	37
7.1.1.4. Mapas SAS	38
7.1.1.5. Caracterización cuantitativa de los mapas SAS	42
7.1.1.6. Activity cliffs profundos y superficiales	45
7.1.1.7. Grado de consenso	46
7.1.1.8. Activity cliffs profundos y superficiales consenso, y activity cliffs aparentes	52
7.1.1.9. Modelos consenso del panorama de actividad: mapas SAS consenso	53
7.1.2. Panorama de actividad y selectividad para 55 derivados de bencimidazol	54
7.1.2.1. Selección de huellas digitales	55
7.1.2.2. Panorama de actividad para el nuevo grupo de 55 moléculas	56
7.1.2.3. Relación entre la actividad contra ambos parásitos y mapas de Diferencia de Actividad Dual (DAD)	60
7.1.2.4. Activity cliffs con mapas DAD	65
7.2. Relaciones cuantitativas estructura–actividad de derivados de bencimidazol con actividad tricomonocida	67
7.2.1. Modelos construidos para selección de tautómeros	69
7.2.2. Modelos de CoMFA y CoMSIA usando confórmeros de mínima energía	72
7.2.3. Modelos construidos con conformaciones selectas empleando similitud 3D	74
7.2.4. Validación	76
7.2.5. Mapas de contornos de CoMFA y CoMSIA	78
7.2.5.1. Mapas de contorno estéricos para el mejor modelo de CoMFA	79
7.2.5.2. Mapas de contorno electrostáticos para el mejor modelo de CoMFA	80
7.2.5.3. Mapas de contorno estéricos para el mejor modelo de CoMSIA	81

7.2.5.4. Mapas de contorno electrostáticos para el mejor modelo de CoMSIA	81
7.2.5.5. Mapas de contorno de donadores y aceptores de hidrógeno para el mejor modelo de CoMSIA	82
7.2.5.6. Mapas de contorno hidrófobos para el mejor modelo de CoMSIA	83
7.3. Actividad antiprotozoaria de los inhibidores de la bomba de protones	84
7.3.1. Inhibidores de la bomba de protones	84
7.3.2. Similitud entre PPIs y algunos bencimidazoles antiprotozoarios	84
7.3.3. Evaluación antiprotozoaria de los PPIs	86
7.4. Síntesis de nuevos derivados de bencimidazol con actividad antiprotozoaria	88
7.4.1. Diseño de las moléculas	88
7.4.2. Síntesis de 1,2-Diaminobencenos (2–17)	90
7.4.2.1. 4-Cloro-1,2-diaminobenceno (2)	90
7.4.2.2. 4,5-Dicloro-1,2-diaminobenceno (3)	91
7.4.2.3. 4-Cloro-N ² -metil-1,2-diaminobenceno (4)	92
7.4.2.4. 5-Cloro-N ¹ -metil-1,2-diaminobenceno (5)	93
7.4.2.5. 3,4-Diaminobenzoato de metilo (6)	93
7.4.2.6. 4,5-Diamino-2-clorobenzoato de metilo (7)	94
7.4.2.7. 4-Amino-3-(metilamino)benzoato de metilo (8)	94
7.4.2.8. 4-Amino-2-cloro-5-(metilamino)benzoato de metilo (9)	95
7.4.2.9. 3-Amino-4-(metilamino)benzoato de metilo (10)	96
7.4.2.10. 5-Amino-2-cloro-4-(metilamino)benzoato de metilo (11)	97
7.4.2.11. 1,2-Diamino-4-etoxibenceno (12)	97
7.4.2.12. 1,2-Diamino-4-Cloro-5-etoxibenceno (13)	98
7.4.2.13. 4-Etoxi-N ² -metil-1,2-diaminobenceno (14)	99
7.4.2.14. 4-Cloro-5-etoxi-N-metil-1,2-diaminobenceno (15)	100
7.4.2.15. 4-Etoxi-N ¹ -metil-1,2-diaminobenceno (16)	101
7.4.2.16. 5-Cloro-4-etoxi-N-metil-1,2-diaminobenceno (17)	101
7.4.3. Síntesis de los 2-Mercaptobencimidazoles (18–34)	102
7.4.4. Síntesis del 1-(2-Cloroetil)-1H-imidazol	103
7.4.5. Síntesis de los compuestos (S1J01–S3J19)	104
7.4.6. Caracterización de compuestos finales	105
7.4.6.1. RMN 1H de los compuestos finales	106
7.4.6.2. RMN ¹³ C de los compuestos finales	111
7.4.6.3. Espectrometría de masas y análisis elemental de los compuestos finales	115
7.4.7. Evaluación de la actividad antiprotozoaria de los compuestos finales	117
8. Conclusiones	120
9. Procedimientos experimentales	122
9.1 Instrumentación	122

9.2. Sistemas cromatográficos	123
9.3. Extracción de los inhibidores de la bomba de protones	123
9.3.1. 5(6)-Metoxi-2-[[4-metoxi-3,5-dimetilpiridin-2-il)metil]sulfinil]-1H-bencimidazol (omeprazol)	124
9.3.2. 2-[[[3-Metil-4-(2,2,2-trifluoroetoxi)piridin-2-il)metil]sulfinil]-1H-bencimidazol (lansoprazol)	124
9.3.3. 2-[[[4-(3-Metoxipropoxi)-3-metilpiridin-2-il)metil]sulfinil]bencimidazol, sal de sodio (rabeprazol sódico)	124
9.3.4. 5(6)-(Difluorometoxi)-2-[[3,4-dimetoxipiridin-2-il)metil]sulfinil]-1H-bencimidazol (pantoprazol)	125
9.4. Procedimientos para la obtención de compuestos finales	125
9.4.1. Procedimiento general de alquilación con 1-(2-cloroetil)-1H-imidazol (S1J01–S1J03, S1J05, S1J06, y S2J08–S3J19)	125
9.4.2. Procedimiento general de alquilación con yoduro de metilo (S1J04 y S1J07)	125
9.4.1.1. 2-[[2-(1H-Imidazol-1-il)etil]sulfanil]-1H-bencimidazol (S1J01)	126
9.4.1.2. 5-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (S1J02)	126
9.4.1.3. 5,6-Dicloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (S1J03)	126
9.4.1.4. 2-[[2-(1H-Imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S1J04)	127
9.4.1.5. 6-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S1J05)	127
9.4.1.6. 5-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S1J06)	128
9.4.1.7. 5,6-Dicloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S1J07)	128
9.4.1.8. 2-[[2-(1H-Imidazol-1-il)etil]sulfanil]-1H-bencimidazol-5-carboxilato de metilo (S2J08)	129
9.4.1.9. 6-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol-5-carboxilato de metilo (S2J09)	129
9.4.1.10. 2-[[2-(1H-Imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-6-carboxilato de metilo (S2J10)	130
9.4.1.11. 5-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-6-carboxilato de metilo (S2J11)	130
9.4.1.12. 2-[[2-(1H-Imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-5-carboxilato de metilo (S2J12)	131
9.4.1.13. 6-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-5-carboxilato de metilo (S2J13)	131
9.4.1.14. 5-Etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (S3J14)	132
9.4.1.15. 6-Cloro-5-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (S3J15)	132
9.4.1.16. 6-Etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S3J16)	133
9.4.1.17. 5-Cloro-6-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S3J17)	133
9.4.1.18. 5-Etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S3J18)	134
9.4.1.19. 6-Cloro-5-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S3J19)	134
9.5. Procedimiento de obtención de 2-mercaptobencimidazoles	135
9.5.1. Procedimiento general (18–34)	135
9.5.1.1. 2-Mercapto-1H-bencimidazol (18)	135

9.5.1.2. 5-Cloro-2-mercapto-1H-bencimidazol (19)	135
9.5.1.3. 5,6-Dicloro-2-mercapto-1H-bencimidazol (20)	135
9.5.1.4. 6-Cloro-2-mercapto-1-metil-1H-bencimidazol (21)	136
9.5.1.5. 5-Cloro-2-mercapto-1-metil-1H-bencimidazol (22)	136
9.5.1.6. 2-Mercapto-1H-bencimidazol-5-carboxilato de metilo (23)	136
9.5.1.7. 6-Cloro-2-mercapto-1H-bencimidazol-5-carboxilato de metilo (24)	136
9.5.1.8. 1-Metil-2-mercapto-1H-bencimidazol-6-carboxilato de metilo (25)	137
9.5.1.9. 5-Cloro-2-mercapto-1-metil-1H-bencimidazol-6-carboxilato de metilo (26)	137
9.5.1.10. 2-Mercapto-1-metil-1H-bencimidazol-5-carboxilato de metilo (27)	137
9.5.1.11. 6-Cloro-2-mercapto-1-metil-1H-bencimidazol-5-carboxilato de metilo (28)	138
9.5.1.12. 5-Etoxi-2-mercapto-1H-bencimidazol (29)	138
9.5.1.13. 6-Cloro-5-etoxi-2-mercapto-1H-bencimidazol (30)	138
9.5.1.14. 6-Etoxi-2-mercapto-1-metil-1H-bencimidazol (31)	138
9.5.1.15. 5-Cloro-6-etoxi-2-mercapto-1-metil-1H-bencimidazol (32)	139
9.5.1.16. 5-Etoxi-2-mercapto-1-metil-1H-bencimidazol (33)	139
9.5.1.17. 6-Cloro-5-etoxi-2-mercapto-1-metil-1H-bencimidazol (34)	139
9.6. Procedimiento de obtención de 1,2-Diaminobencenos (o-fenilendiaminas)	140
9.6.1. Procedimiento general de hidrogenación catalítica	140
9.6.2. 3,4-Diaminobenzoato de metilo (6)	140
9.7. Preparación del 1-(2-Cloroetil)-1H-imidazol	140
9.7.1. 1-(2-Cloroetil)-1H-imidazol	140
9.8. Síntesis de precursores	141
9.8.1. Procedimientos generales	141
9.8.1.1. Procedimiento general de acetilación	141
9.8.1.2. Procedimiento general para hidrólisis de acetamidas en ácido sulfúrico concentrado	141
9.8.1.3. Procedimiento general para hidrólisis de acetamidas en ácido sulfúrico y metanol	141
9.8.1.4. Procedimiento general para hidrólisis de acetamidas en medio básico	142
9.8.1.5. Procedimiento general de N-metilación de acetamidas	142
9.8.1.6. Procedimiento general para la preparación de ésteres con sulfato de dimetilo	142
9.8.1.7. Procedimiento general para realizar la esterificación de Fischer	142
9.8.1.8. Procedimiento general de alquilación de fenoles	143
9.8.2. Preparación de intermediarios	143
9.8.2.1. N-(3,4-Diclorofenil)acetamida (Int-3)	143
9.8.2.2. N-(4,5-Dicloro-2-nitrofenil)acetamida (Int-4)	143
9.8.2.3. 4,5-Dicloro-2-nitroanilina (Int-5)	144
9.8.2.4. 2,4-Dicloronitrobenzeno (Int-7)	144

9.8.2.5. 5-Cloro-N-metil-2-nitroanilina (Int-8)	144
9.8.2.6. N-(4-Cloro-2-nitrofenil)acetamida (Int-9)	145
9.8.2.7. N-(4-Cloro-2-nitrofenil)-N-metilacetamida (Int-10)	145
9.8.2.8. 4-Cloro-N-metil-2-nitroanilina (Int-11)	145
9.8.2.9. 3-Metoxi-4-nitrobenzoato de metilo (Int-15)	145
9.8.2.10. Ácido 3-metoxi-4-nitrobenzoico (Int-16)	146
9.8.2.11. Ácido 3-(metilamino)-4-nitrobenzoico (Int-17)	146
9.8.2.12. 3-(Metilamino)-4-nitrobenzoato de metilo (Int-18)	146
9.8.2.13. 2-Cloro-5-nitrobenzoato de metilo (Int-20)	147
9.8.2.14. 5-Amino-2-clorobenzoato de metilo (Int-21)	147
9.8.2.15. 5-(Acetamino)-2-clorobenzoato de metilo (Int-22)	147
9.8.2.16. 5-(Acetamino)-2-cloro-4-nitrobenzoato de metilo (Int-23a)	147
9.8.2.17. 5-[Acetil(metil)amino]-2-cloro-4-nitrobenzoato de metilo (Int-24)	148
9.8.2.18. 2-Cloro-5-(metilamino)-4-nitrobenzoato de metilo (Int-25)	149
9.8.2.19. 4-Cloro-3-nitrobenzoato de metilo (Int-27)	149
9.8.2.20. 4-(Metilamino)-3-nitrobenzoato de metilo (Int-28)	149
9.8.2.21. 4-[Acetil(metil)amino]-2-cloro-5-nitrobenzoato de metilo (Int-30)	150
9.8.2.22. 2-Cloro-4-(metilamino)-5-nitrobenzoato de metilo (Int-31)	150
9.8.2.23. 1-Etoxi-4-nitrobenceno (Int-33)	150
9.8.2.24. 4-Etoxianilina (Int-34)	150
9.8.2.25. N-4-(Etoxifenil)acetamida (Int-35)	151
9.8.2.26. N-(4-Etoxi-2-nitrofenil)acetamida (Int-36)	151
9.8.2.27. 4-Etoxi-2-nitroanilina (Int-37)	151
9.8.2.28. 4-Cloro-5-etoxi-2-nitroanilina (Int-38)	151
9.8.2.29. 3-Cloro-4-nitrofenol (Int-40)	152
9.8.2.30. 2-Cloro-4-etoxi-1-nitrobenceno (Int-41)	152
9.8.2.31. 5-Etoxi-N-metil-2-nitroanilina (Int-42)	152
9.8.2.32. N-(4-Cloro-5-etoxi-2-nitrofenil)acetamida (Int-43)	153
9.8.2.33. N-(4-Cloro-5-etoxi-2-nitrofenil)-N-metilacetamida (Int-44)	153
9.8.2.34. 4-Cloro-5-etoxi-N-metil-2-nitroanilina (Int-45)	153
9.8.2.35. N-(4-Etoxi-2-nitrofenil)-N-metilacetamida (Int-46)	153
9.8.2.36. 4-Etoxi-N-metil-2-nitroanilina (Int-47)	154
9.8.2.37. 2-Cloro-1-etoxi-4-nitrobenceno (Int-49)	154
9.8.2.38. 3-Cloro-4-etoxianilina (Int-50)	154
9.8.2.39. N-(3-Cloro-4-etoxifenil)acetamida (Int-51)	154
9.8.2.40. N-(5-Cloro-4-etoxi-2-nitrofenil)acetamida (Int-52)	155
9.8.2.41. 5-Cloro-4-etoxi-2-nitroanilina (Int-53)	155

9.8.2.42. <i>N</i> -(5-Cloro-4-etoxi-2-nitrofenil)- <i>N</i> -metilacetamida (Int-54)	155
9.8.2.43. 5-Cloro-4-etoxi- <i>N</i> -metil-2-nitroanilina (Int-55)	156
10. Referencias	157
Apéndice I: Espectros	165
Apéndice II: Publicaciones	223

ÍNDICE DE TABLAS

	Pag.
Tabla 1 Moléculas diseñadas	27
Tabla 2 Estructura, actividad tricomonocida y giardicida de derivados de bencimidazol usados en el estudio de <i>panorama de actividad</i>	33
Tabla 3 Representaciones moleculares calculadas	35
Tabla 4 Correlación entre diferentes representaciones de similitud molecular	37
Tabla 5 Distribución de datos en cada región de los mapas SAS	42
Tabla 6 Número total de <i>activity cliffs superficiales</i> en diferentes representaciones moleculares	45
Tabla 7 Número de pares consenso en las cuatro regiones de los mapas SAS para las representaciones moleculares: Radial, MACCS, TGD, piDAPH3 y Similitud en Propiedades (SP).	47
Tabla 8 Grado de consenso (Doc) en las cuatro regiones de los mapas SAS para las representaciones moleculares: Radial, MACCS, TGD, piDAPH3 y Similitud en Propiedades (SP).	48
Tabla 9 Ejemplos de pares consenso en los mapas SAS	49
Tabla 10 Matriz de correlación para las huellas digitales	55
Tabla 11 Distribución de puntos en las diferentes regiones de los mapas SAS	58
Tabla 12 Estructura y actividad tricomonocida de los derivados de bencimidazol usados en los estudios QSAR-3D	67
Tabla 13 Resumen de resultados para los modelos de CoMFA 1–5	70
Tabla 14 Valores de pCl ₅₀ experimentales y calculados con modelo 2 para moléculas con tautómeros	72
Tabla 15 Resumen de resultados para los modelos de CoMFA 6–12	73
Tabla 16 Resumen de resultados para los modelos de CoMFA 13–17	75
Tabla 17 Predicciones para el grupo de prueba usando los modelos 6, 12, 13 y 17	78
Tabla 18 Actividad in vitro de los PPIs comerciales contra <i>T. vaginalis</i> , <i>G. intestinalis</i> y <i>E. histolytica</i>	86
Tabla 19 Actividad calculada para las moléculas diseñadas	90
Tabla 20 Rendimiento y punto de fusión de los 2-mercaptobencimidazoles sintetizados	102
Tabla 21 Rendimiento y punto de fusión para los compuestos finales	105
Tabla 22 Resumen de las señales de RMN ¹ H encontradas a campo alto para los compuestos finales	107
Tabla 23 Resumen de las señales de RMN ¹ H encontradas a campo bajo para los compuestos finales	109
Tabla 24 Desplazamiento químico para los carbonos alifáticos de los compuestos finales	112
Tabla 25 Desplazamiento químico para los carbonos aromáticos de los compuestos finales	113

Tabla 26	Análisis elemental por espectrometría de masas de alta resolución	117
Tabla 27	Actividad antiprotozoaria de los compuestos sintetizados	118
Tabla 28	Predicciones para las moléculas sintetizadas usando los mejores modelos QSAR-3D (6, 12, 13 y 17) y su comparación con los resultados experimentales	119
Tabla 29	Sistemas cromatográficos usados en este trabajo	123

ÍNDICE DE FIGURAS

		Pag.
Figura 1	Estructura del metronidazol.	3
Figura 2	Estructura del albendazol y mebendazol.	3
Figura 3	Derivados de bencimidazol con actividad antiprotozoaria.	4
Figura 4	Derivados de bencimidazol 1-metilados.	4
Figura 5	Derivados de benzoxazol.	5
Figura 6	Derivados de 2-(trifluorometil) bencimidazol con actividad antiprotozoaria.	6
Figura 7	Derivados de 2-(trifluorometil) y 2-(tentafluoroetil)-1 <i>H</i> -bencimidazol con actividad antiprotozoaria.	6
Figura 8	Derivados de bencimidazol con actividad antiprotozoaria sustituidos en posiciones 2, 4(7) y 6(5).	7
Figura 9	Derivados de 2-(metiltio)-1 <i>H</i> -bencimidazol con diferentes carboxamidas en posición 5(6).	7
Figura 10	Derivados de 1-metil-1 <i>H</i> -bencimidazol con diferentes sustituciones en posición 2.	8
Figura 11	Clasificación de algunas metodologías usadas en DiFAC.	10
Figura 12	Mapa de contornos para el mejor modelo CoMFA construido con derivados de bencimidazol activos contra <i>E. histolytica</i> .	11
Figura 13	Mapa de contornos para el mejor modelo CoMFA construido con derivados de bencimidazol activos contra <i>G. intestinalis</i> .	12
Figura 14	Forma esquemática un mapa SAS dividido en cuatro regiones (I–IV).	13
Figura 15	Funciones de distribución acumulada para distintas representaciones de similitud estructural en 2D y 3D.	36
Figura 16	Mapas SAS para <i>T. vaginalis</i> con diferentes representaciones moleculares.	40
Figura 17	Mapas SAS para <i>G. intestinalis</i> con diferentes representaciones moleculares.	41
Figura 18	Ejemplos de gráficos circulares para <i>T. vaginalis</i> .	44
Figura 19	Compuestos incluidos en algunos pares consenso representativos para <i>T. vaginalis</i> y <i>G. intestinalis</i> .	50
Figura 20	Mapas SAS consenso que muestran puntos con tamaño proporcional a la desviación estándar.	54
Figura 21	Funciones de distribución acumulada para distintas representaciones de similitud estructural en 2D y 3D.	56
Figura 22	Mapas SAS para <i>T. vaginalis</i> y <i>G. intestinalis</i> con tres diferentes representaciones moleculares	57
Figura 23	Mapa SAS representativo para <i>T. vaginalis</i> obtenido con la huella digital MACCS.	59
Figura 24	Correlación entre los valores de pCl ₅₀ para <i>T. vaginalis</i> y <i>G. intestinalis</i> .	61
Figura 25	Forma general de un mapa DAD donde se muestran las cuatro zonas principales Z1–Z4.	61
Figura 26	Mapa DAD para la base de datos de derivados de bencimidazol con actividad contra <i>T. vaginalis</i> y <i>G. intestinalis</i> .	62
Figura 27	Estructuras químicas de los compuestos incluidos en los pares selectos con los valores de Cl ₅₀ para <i>T. vaginalis</i> (Tv) y <i>Giardia intestinalis</i> (Gi).	64
Figura 28	Mapa DAD con pares coloreados por similitud estructural.	65
Figura 29	Mapa DAD que muestra 925 pares con alta similitud estructural.	66
Figura 30	Formas tautoméricas consideradas en este trabajo y su identificación basada en su patrón de sustitución.	69
Figura 31	Orientación de los sustituyentes en las posiciones 5 y 6 para los modelos 1–5.	71

Figura 32	Moléculas incluidas en el grupo de trabajo usando (a) las conformaciones de mínima energía y (b) Las conformaciones seleccionadas por superposición con 9 , 10 y 45 .	75
Figura 33	Validación del modelo de CoMFA 6.	77
Figura 34	Mapas de contornos para los mejores modelos (6 y 12).	79
Figura 35	Algunos ejemplos de derivados de bencimidazol con actividad antiprotozoaria y los PPIs comerciales.	84
Figura 36	Sobreposición de los compuestos de referencia 1–4 y los PPIs.	85
Figura 37	Potencial electrostático molecular para los derivados 1–4 y los PPIs.	86
Figura 38	Conclusiones más importantes sobre relaciones estructura-actividad para bencimidazoles evaluados contra <i>T. vaginalis</i> , <i>G. intestinalis</i> y <i>E. histolytica</i> .	88
Figura 39	Consideraciones para el diseño del compuesto S1J01 .	89
Figura 40	Relación entre la estructura general de los compuestos finales y las señales de RMN ¹ H observadas.	106
Figura 41	Espectro de RMN ¹ H del compuesto S2J10 (A) y producto secundario (B).	110
Figura 42	Relación entre la estructura general de los compuestos finales y las señales de RMN ¹³ C observadas.	111
Figura 43	Efecto de la tautomería en los espectros de RMN ¹³ C.	114
Figura 44	Fragmentos propuestos, para algunos compuestos finales representativos, mediante la interpretación de sus espectros de masas.	116

ÍNDICE DE ESQUEMAS

		Pag.
Esquema 1	Reactivos y condiciones para la síntesis de los compuestos S1J01–S3J19	28
Esquema 2	Preparación del reactivo 1-(2-Cloroetil)-1 <i>H</i> -imidazol.	29
Esquema 3	Síntesis de los 1,2-diaminobencenos 2–5 a partir de compuestos comerciales.	29
Esquema 4	Síntesis de los 1,2-diaminobencenos 6–11 a partir de compuestos comerciales.	30
Esquema 5	Síntesis de los 1,2-diaminobencenos 12–17 a partir de compuestos comerciales.	31
Esquema 6	Síntesis del compuesto 2 .	91
Esquema 7	Síntesis del compuesto 3 .	91
Esquema 8	Síntesis del compuesto 4 .	92
Esquema 9	Síntesis del compuesto 5 .	93
Esquema 10	Síntesis del compuesto 6 .	94
Esquema 11	Síntesis del compuesto 7 .	94
Esquema 12	Síntesis del compuesto 8 .	94
Esquema 13	Síntesis del compuesto 9 .	96
Esquema 14	Síntesis del compuesto 10 .	97
Esquema 15	Síntesis del compuesto 11 .	97
Esquema 16	Síntesis del compuesto 12 .	98
Esquema 17	Síntesis del compuesto 13 .	99
Esquema 18	Síntesis del compuesto 14 .	99
Esquema 19	Síntesis del compuesto 15 .	100
Esquema 20	Síntesis del compuesto 16 .	101
Esquema 21	Síntesis del compuesto 17 .	101
Esquema 22	Rutas planteadas para la preparación del 1-(2-cloroetil)-1 <i>H</i> -imidazol.	103

ABREVIATURAS Y ACRÓNIMOS

SAR	Del inglés: Structure Activity Relationships
-----	--

MOE	Del inglés: Molecular Operating Environment
SAS	Del inglés: Structure Activity Similarity
DAD	Del inglés: Dual Activity Difference
DoC	Del inglés: Degree of Consensus
Ph4	Farmacóforo
DSA	División de Similitud en Actividad
CoMFA	Del inglés: Comparative Molecular Field Analysis
CoMSIA	Del inglés: Comparative Molecular Similarity Indices Analysis
QSAR	Del inglés: Quantitative Structure Activity Relationships
PLS	Del inglés: Partial Least Squares
ROCS	Del inglés: Rapid Overlay of Chemical Structures
LOO	Del inglés: Leave One Out
PPIs	Del inglés: Proton Pump Inhibitors
ccf	Cromatografía en capa fina
pf	Punto de fusión
dec	Descomposición
s	Singulete
d	Doblete
dd	Doble de dobles
q	Del inglés: quartet
bs	Del inglés: broad signal
eq	Equivalentes

RESUMEN

Los derivados de bencimidazol son compuestos versátiles con un amplio rango de actividades biológicas, entre ellas se conocen la antihelmíntica, antiprotozoaria, antiúlceras, antifúngica, antitumoral, etc. A pesar de que el uso de métodos computacionales basados en quimioinformática y modelado molecular permite el estudio y diseño de nuevos fármacos de una manera eficiente, poco se ha reportado acerca de estos nuevos métodos en el diseño de derivados de bencimidazol antiparasitarios. En esta tesis se describen estudios de relaciones estructura-actividad, diseño racional y computacional, síntesis y evaluación antiprotozoaria de derivados de bencimidazol. El trabajo realizado se presenta en cuatro partes.

La primera parte describe estudios sobre el panorama de actividad de derivados de bencimidazol con actividad tricomonocida y giardicida. Se analizó una base de datos, evaluada contra *Trichomonas vaginalis* y *Giardia intestinalis*, empleando mapas de similitud estructura-actividad (SAS). Los resultados de este estudio mostraron que existe un SAR heterogéneo y algunas características estructurales importantes para diseñar nuevos agentes antiparasitarios bencimidazólicos potentes. Posteriormente, el estudio se extendió mediante el uso de un nuevo método, desarrollado en la presente tesis, el cual se denominó mapas de diferencia de actividad dual (DAD). Los resultados mostraron que existe relación entre las actividades giardicida y tricomonocida; además se encontraron algunas características estructurales que dirigen a la optimización de la actividad contra ambos protozoarios.

La segunda parte de este trabajo incluye estudios de relaciones cuantitativas estructura-actividad (QSAR) para derivados de bencimidazol con actividad tricomonocida, empleando las metodologías conocidas como análisis comparativo de campos moleculares (CoMFA) y análisis comparativo de índices de similitud (CoMSIA). Los resultados obtenidos muestran una alta correlación entre la estructura y la actividad. Adicionalmente, los mapas de contornos generados indican algunas características importantes que pueden ser usadas para el diseño de nuevos compuestos con potente actividad tricomonocida.

En la tercera parte de este trabajo, basados en los estudios SAR previos y en algunos estudios de modelado molecular adicionales, se descubrió la actividad

antiprotozoaria de algunos derivados de bencimidazol conocidos como inhibidores de la bomba de protones (PPIs). Los PPIs se obtuvieron de sus presentaciones comerciales y se sometieron a evaluaciones in vitro contra *T. vaginalis*, *G. intestinalis* y *Entamoeba histolytica*. Los ensayos biológicos revelaron una potente actividad antiprotozoaria de los PPIs contra los tres protozoarios.

Finalmente, la última parte de este trabajo describe el diseño y síntesis de 19 nuevos derivados de 2-{{2-(1*H*-imidazol-1-il)etil}sulfanil}-1*H*-bencimidazol. Las evaluaciones in vitro contra *T. vaginalis*, *G. intestinalis* y *E. histolytica* mostraron que los compuestos tienen actividad potente. Además se observó que las predicciones de actividad tricomonocida derivadas de los modelos QSAR son acertadas en la mayoría de los casos.

ABSTRACT

Benzimidazole derivatives are versatile compounds with a wide range of biological activities; among these are known anthelmintic, antiprotozoal, antiulcer, antifungal, antitumoral, antiviral, etc. Although the application of computational methods based on chemoinformatics and molecular modeling allow the study and design of new drugs efficiently, little is reported about the use of these novel methods in the design of antiparasitic benzimidazole derivatives. This thesis deals with the structure–activity relationships studies, rational and computational design, synthesis and antiprotozoal evaluation of benzimidazole derivatives. The work done is presented in four parts.

The first part describes activity landscape studies of benzimidazole derivatives with trichomonocidal and giardicidal activity. A benzimidazole dataset, tested against *Trichomonas vaginalis* and *Giardia intestinalis*, was analyzed using Structure Activity Similarity (SAS) maps. Results of this study showed a heterogeneous SAR, also, some structural features that led to the design of potent antiparasitic compounds. This study was extended by using a new approach proposed in this thesis, namely Dual Activity Difference (DAD) maps. A relationship between giardicidal and trichomonocidal activities was found, moreover, some structural features that led to dual antiprotozoal activity optimization.

The second part of this work involves Quantitative Structure-Activity Relationships (QSAR) studies for a series of benzimidazole derivatives, with trichomonocidal activity, employing Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA) methodologies. Results obtained show a high correlation between structure and the activity. Additionally, the contour maps generated indicate some important features useful for the design of new compounds with potent trichomonocidal activity.

In the third part of this work, based on previous SAR studies and some additional molecular modeling studies, the antiprotozoal activity of benzimidazole derivatives known as proton pump inhibitors (PPIs) was discovered. PPIs obtained from commercial presentations were tested in vitro against *T. vaginalis*, *G. intestinalis* and *Entamoeba histolytica*. Biological assays revealed a strong antiparasitic activity of PPIs against the three protozoa tested.

Finally, the last part of this work describes the design and synthesis of 19 new 2-[[2-(1*H*-imidazol-1-yl)ethyl]sulfanyl]-1*H*-benzimidazole derivatives. The evaluation in vitro of these compounds against *T. vaginalis*, *G. intestinalis* and *E. histolytica* showed a potent activity. Furthermore, it was observed that the trichomonocidal activity of the new set of compounds was predicted correctly by QSAR models in most of the cases.

1. INTRODUCCIÓN

Las enfermedades parasitarias causadas por helmintos y protozoarios aun representan un problema de salud pública a escala mundial. Según datos de la Organización Mundial de la Salud (OMS), más de la mitad de la población mundial se encuentra infectada por algún tipo de parásito. Entre las enfermedades parasitarias con mayor incidencia se encuentran aquellas producidas por los protozoarios *Trichomonas vaginalis*, *Giardia intestinalis* y *Entamoeba histolytica*; que en conjunto, afectan cada año a más de 500 millones de habitantes. Estas enfermedades generalmente no son causa de urgencias médicas; sin embargo, producen graves deterioros a la salud a largo plazo.

En el control de las enfermedades producidas por protozoarios, la prevención y la quimioterapia juegan un papel muy importante. El metronidazol, 2-(2-metil-5-nitro-1H-imidazol-1-il)etanol, es el fármaco de elección desde hace más de 40 años contra la tricomonosis, giardosis y amebiosis. Sin embargo, debido a sus efectos adversos y a la aparición de cepas resistentes, existe la necesidad de contar con nuevos agentes terapéuticos.

Durante los últimos años, en nuestro grupo de investigación se ha trabajado en la síntesis de derivados de bencimidazol como agentes antiprotozoarios. Los compuestos obtenidos, hasta la fecha, constituyen una base de datos útil para estudiar las relaciones estructura-actividad y diseñar nuevos compuestos activos empleando quimioinformática y modelado molecular; técnicas que en la actualidad son de gran importancia en el estudio y diseño de fármacos.

Considerando la problemática de las enfermedades producidas por protozoarios, en este trabajo se usaron algunas herramientas computacionales para estudiar las relaciones estructura-actividad de los derivados de bencimidazol con actividad antiprotozoaria y se diseñaron nuevas moléculas para su síntesis y evaluación como antiprotozoarios.

2. ANTECEDENTES

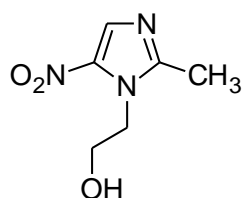
2.1. Enfermedades parasitarias

Las enfermedades parasitarias o parasitosis son un grave problema de salud pública a escala mundial. En nuestro país se reportan entre las principales causas de morbilidad [1]. Dentro de las enfermedades parasitarias con mayor prevalencia se encuentran las infecciones causadas por protozoarios y helmintos. Estas enfermedades pueden adquirirse a través de los alimentos o del agua contaminada (como la fasciolosis o la teniosis), por la picadura de un insecto (como la malaria o la enfermedad del sueño) o por contacto sexual (como la tricomonosis).

Las infecciones más comunes son producidas por protozoarios que invaden las mucosas, afectando a más de un billón de personas cada año. Entre este tipo de enfermedades (protozoosis) se pueden mencionar a la tricomonosis, giardiosis y amebiosis, causadas por *Trichomonas vaginalis*, *Giardia intestinalis* y *Entamoeba histolytica*, respectivamente [2]. De acuerdo a la OMS, las infecciones del tracto genitourinario causadas por *T. vaginalis* se estiman en 170 millones de nuevos casos anuales. Esta parasitosis afecta principalmente a mujeres en las que frecuentemente se presenta inflamación del tracto genitourinario. La gravedad de la infección se presenta porque se ha asociado con esterilidad, cáncer cervical, predisposición a la infección con VIH, labor de pretérmino en mujeres embarazadas y bajo peso en recién nacidos [3–6]. Por otra parte, la enfermedad intestinal conocida como giardiosis afecta a 280 millones de personas cada año. Esta infección causa diarrea y síntomas asociados; además tiene una incidencia muy elevada en niños en los que causa lento desarrollo corporal e intelectual [2–3]. Otra de las enfermedades intestinales importantes es la amebiosis. La OMS estima 50 millones de casos y 100,000 muertes anuales asociadas con esta enfermedad. Los principales síntomas son diarrea y disentería. Este último protozoario es de especial importancia ya que puede penetrar la mucosa intestinal y migrar a otros órganos, produciendo daños que pueden llegar a ser mortales [2, 4].

El metronidazol (MTZ) es el fármaco de elección para el tratamiento de la tricomonosis, giardiosis y amebiosis desde hace más de 40 años (Figura 1). Sin

embargo, dado los efectos adversos ligados a este fármaco y a la aparición de cepas resistentes, existe la necesidad de contar con nuevos agentes terapéuticos [2, 5].

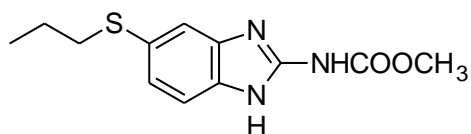


2-(2-Metil-5-nitro-1*H*-imidazol-1-il)etanol

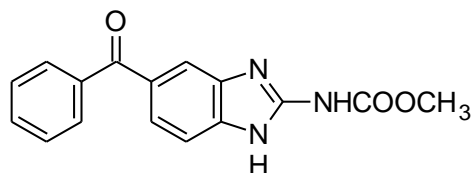
Figura 1. Estructura del metronidazol.

2.2. Bencimidazoles como agentes antiprotozoarios

A finales de la década de los 80s y a principios de los 90s se realizaron estudios concernientes a la evaluación de la actividad antiprotozoaria del albendazol (ABZ) y mebendazol (MBZ). Éstos fueron los primeros reportes sobre la actividad giardicida y tricomonocida para derivados de bencimidazol que hasta la fecha son usados como antihelmínticos (Figura 2) [7–10].



Albendazol



Mebendazol

Figura 2. Estructura del albendazol y mebendazol.

Este descubrimiento estimuló el interés por sintetizar nuevos derivados de bencimidazol y evaluarlos como antiprotozoarios. Entre los primeros trabajos se encuentra la síntesis y evaluación de algunos derivados sustituidos en posición 2 del núcleo de bencimidazol con hidrógeno, metilo, amino, mercapto, metiltio y carbamato de metilo; además con sustituciones que incluyen cloro e hidrógeno en posiciones 5 y 6 (Figura 3) [11].

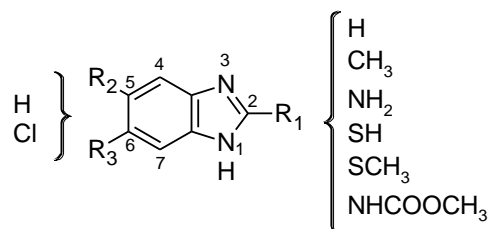


Figura 3. Derivados de bencimidazol con actividad antiprotozoaria.

Los nuevos compuestos sintetizados mostraron, en general, muy buena actividad giardicida (0.005–1.902 μM , mediana 0.080 μM) y amebicida (0.005–0.393 μM , mediana 0.090 μM); en la mayoría de los casos, fueron más activos que el fármaco de referencia metronidazol. En ese trabajo se determinó que el bencimidazol no sustituido y con sustituciones discretas (como se muestra en la Figura 3) tiene potente actividad antiprotozoaria. Adicionalmente, se sabe que los bencimidazoles sustituidos en posición 2 con el grupo carbamato de metilo tienen la capacidad de inhibir la polimerización de la β -tubulina [12], siendo este el mecanismo por el cual ejercen su actividad antihelmíntica; sin embargo, ensayos de estos compuestos muestran que sólo los derivados 2-metilcarbamatos afectan la polimerización de la tubulina. Considerando esta observación, se concluyó que los nuevos bencimidazoles sustituidos con un grupo distinto al carbamato de metilo ejercen su actividad antiprotozoaria por un mecanismo de acción diferente a la inhibición de la polimerización de la β -tubulina.

Con el fin de conocer más acerca de los requerimientos estructurales para la actividad antiprotozoaria de los derivados de bencimidazol, en particular el papel del hidrógeno en posición 1, se prepararon derivados 1-metilados (Figura 4) que mostraron buena actividad giardicida en muchos casos (0.008–143.45 μM , mediana 0.127 μM).

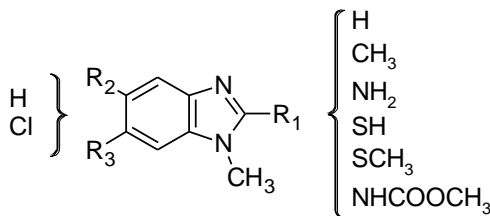


Figura 4. Derivados de bencimidazol 1-metilados.

En estos compuestos se observó que la actividad giardicida es superior a la de sus análogos con hidrógeno en posición 1 en la mayoría de los casos. Estudios de microscopía electrónica de barrido y transmisión mostraron evidencia de un mecanismo de acción diferente en estos compuestos con respecto a los derivados de bencimidazol con hidrógeno en posición 1. Por otra parte, se encontró buena actividad amebicida en la mayoría de los casos (0.005–71.62 μM , mediana 0.239 μM) y actividades tricomonocidas que van desde buenas a pobres (0.148–142.3 μM , mediana 3.7945 μM). Particularmente, para esta serie, los derivados con hidrógeno en posición 2 tuvieron las actividades más bajas ante los tres protozoarios [13].

Posteriormente se sintetizaron y evaluaron algunos derivados de benzoxazol contra *G. intestinalis* (Figura 5). Los resultados mostraron actividades en un rango de 0.357–29.5 μM (mediana 0.3681 μM) [14]. Aunque estos derivados tienen una actividad importante, de manera general, son menos activos que los análogos derivados de 1*H*-bencimidazol y del 1-metilbencimidazol.

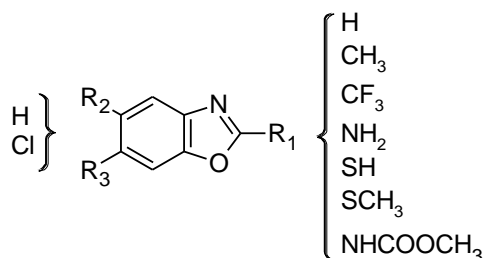


Figura 5. Derivados de benzoxazol.

Una nueva generación de derivados de bencimidazol surgió de la modificación por trifluorometilo en posición 2 y variaciones en los sustituyentes en posiciones 1, 5 y 6 (divididos en series A–C basados en que son producto de diferentes publicaciones) (Figura 6). Los compuestos derivados de 2-(trifluorometil)bencimidazol de la serie A mostraron, en general, buena actividad antiprotozoaria contra *G. intestinalis* (0.042–1.282 μM , mediana 0.107 μM) y *E. histolytica* (0.008–0.069 μM , mediana 0.033 μM) [15]. Como era de esperarse, estudios adicionales demostraron que los derivados de 2-(trifluorometil)bencimidazol tampoco inhibieron la polimerización de la tubulina [15].

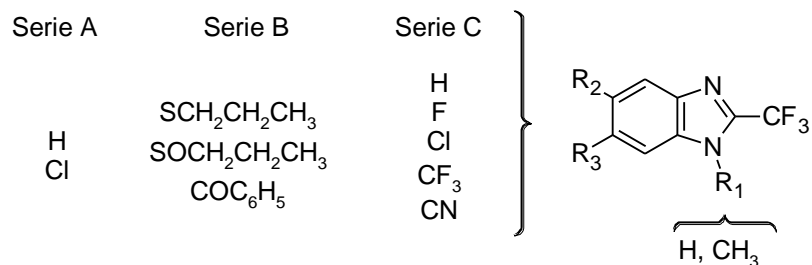


Figura 6. Derivados de 2-trifluorometil bencimidazol con actividad antiprotozoaria.

Posteriormente, en la serie B se diseñaron análogos de albendazol y mebendazol con hidrógeno y metilo en la posición 1 y trifluorometilo en la posición 2 [16]. Los resultados mostraron que la actividad giardicida disminuye con respecto a los derivados de la serie A (1.098–20.89 μ M, mediana 2.31 μ M). Además se evaluó la actividad tricomonocida la serie B y resultaron compuestos con actividad pobre en la mayoría de los casos (0.0587–29.59 μ M, mediana 2.78 μ M). La serie C se basó en que los compuestos clorados de la serie A presentaron buenas actividades, por lo que en esta serie se incluyeron sustituciones en las posiciones 5 y 6 por grupos que se consideran bioisostéricos. Los compuestos de esta serie mostraron actividad pobre, en la mayoría de los casos, contra los protozoarios *T. vaginalis* (0.451–5.32 μ M, mediana 3.134 μ M) y *G. intestinalis* (0.107–8.50 μ M, mediana 1.144 μ M) [17].

Posteriormente, otro grupo de investigación diseñó una nueva serie de derivados de 2-trifluorometil y 2-pentafluoroetilbencimidazol con sustituciones en las posiciones 4–7 del bencimidazol (Figura 7). Estos compuestos mostraron, en la mayoría de los casos, buena actividad giardicida (0.023–0.64 μ M, mediana 0.14 μ M), tricomonocida (0.002–10.1 μ M, mediana 0.27 μ M) y amebicida (0.05–13.80 μ M, mediana 0.3 μ M) [18].

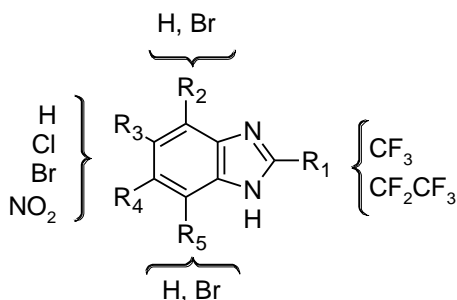


Figura 7. Derivados de 2-(trifluorometil) y 2-(pentafluoroetil)bencimidazol con actividad antiprotozoaria.

En el mismo grupo de investigación se sintetizaron compuestos con nuevas sustituciones en posiciones 2, 4(7) y 6(5) del bencimidazol (Figura 8). Los nuevos derivados mostraron, en la mayoría de los casos, una excelente actividad tricomonocida (0.003–0.692 μM , mediana 0.055 μM) y giardicida (0.019–0.126 μM , mediana 0.038 μM) [19].

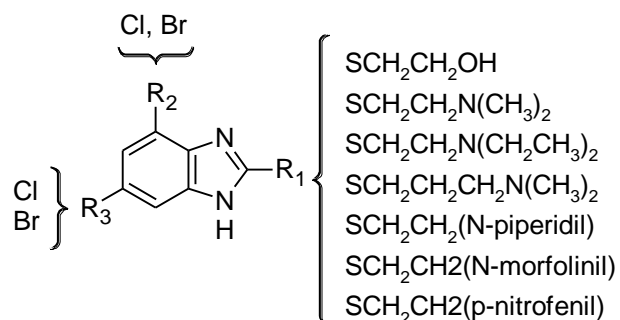


Figura 8. Derivados de bencimidazol con actividad antiprotozoaria sustituidos en posiciones 2, 4(7) y 6(5).

Para enriquecer la base de datos con diferentes sustituyentes en posición 5(6), en nuestro grupo de investigación se sintetizaron y evaluaron algunas carboxamidas bencimidazólicas (Figura 9), las cuales mostraron buena actividad tricomonocida (0.029–0.897, mediana 0.162) y giardicida (0.018–0.345, mediana 0.072) [20].

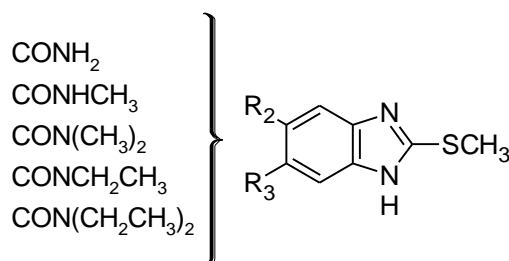


Figura 9. Derivados de 2-metilbencimidazol con diferentes carboxamidas en posición 5(6).

Uno de los trabajos más recientes publicados por nuestro grupo de investigación, trata sobre la síntesis de nuevos derivados de 1-metilbencimidazol que incluyen amidas, ésteres, alcoholes y cetonas en posición 2 (Figura 10). Estos compuestos mostraron una buena actividad tricomonocida (0.0189–0.4305 μM , mediana 0.1206 μM) y giardicida (0.0203–0.5117 μM , mediana 0.0699 μM). Particularmente, los derivados de tipo éster etílico en posición 2 fueron los más potentes [21].

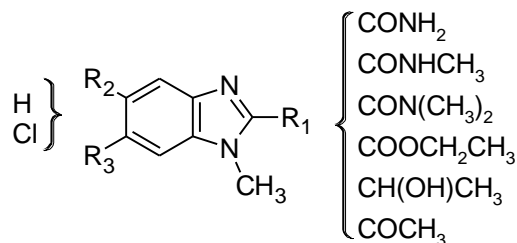


Figura 10. Derivados de 1-metilbencimidazol con diferentes sustituciones en posición 2.

La información generada hasta ahora sobre los derivados de bencimidazol constituye una base de datos importante para el estudio de las relaciones estructura-actividad. Para esto, los métodos computacionales son herramientas muy valiosas que pueden proporcionar información para el diseño de nuevas moléculas de una manera eficiente [22–25].

2.3. *Diseño de fármacos asistido por computadora y su aplicación en la búsqueda de agentes antiprotozoarios*

2.3.1. *Diseño de fármacos asistido por computadora*

El diseño de fármacos asistido por computadora (DiFAC) es un área del diseño racional de fármacos que consiste en aplicar algún procedimiento realizado por una computadora para relacionar la actividad biológica de un compuesto con su estructura química [22, 24, 26]. El DiFAC puede ser usado como herramienta para el descubrimiento y optimización de moléculas activas. Las estrategias de DiFAC se pueden dividir en dos grupos que dependerán de la información disponible para realizar los estudios: estos grupos son el diseño basado en el ligando y diseño basado en el receptor (Figura 11).

Cuando conocemos la estructura tridimensional de un blanco farmacológico potencial, el diseño puede ser dirigido mediante la construcción de nuevas estructuras con sustituyentes que tengan posibilidad de generar interacciones razonables con el sitio activo (diseño *de novo*) [27]. Otra de las metodologías usadas para el diseño basado en el receptor es el acoplamiento molecular (*Docking*); útil en la búsqueda de fármacos a partir de bases de datos virtuales de estructuras químicas [28–29]. Ambas

técnicas pueden combinarse para generar moléculas de una manera racional y proponer un modo de unión probable en el receptor mediante cálculos de energía libre.

Si no conocemos la estructura del receptor, existen técnicas basadas en la estructura del ligado. Para aplicar estas técnicas se requieren moléculas con actividad medida experimentalmente. Algunas de estas técnicas son basadas en farmacóforos [30], similitud molecular [31], relaciones cuantitativas estructura-actividad [32–33], entre otras.

El mejor de los casos es cuando contamos con el receptor o receptores involucrados y moléculas evaluadas experimentalmente, ya que es posible conjuntar metodologías y relacionar los resultados obtenidos. Un ejemplo es la combinación de las relaciones cuantitativas estructura-actividad en tres dimensiones (QSAR-3D) con el acoplamiento molecular, donde ambas técnicas se complementan para formar modelos muy informativos y con alto poder predictivo [34]. En contraparte, cuando no se conoce el receptor ni moléculas activas, la química combinatoria y la química orientada a la diversidad juegan un papel muy importante para generar bibliotecas de compuestos que cubran un amplio espacio químico y puedan usarse para cribado de alto rendimiento (*high-throughput screening*) [35–36]. Así mismo, en el estudio de bibliotecas derivadas de la química combinatoria se ha demostrado la utilidad de los métodos de quimioinformática [24, 37–40].

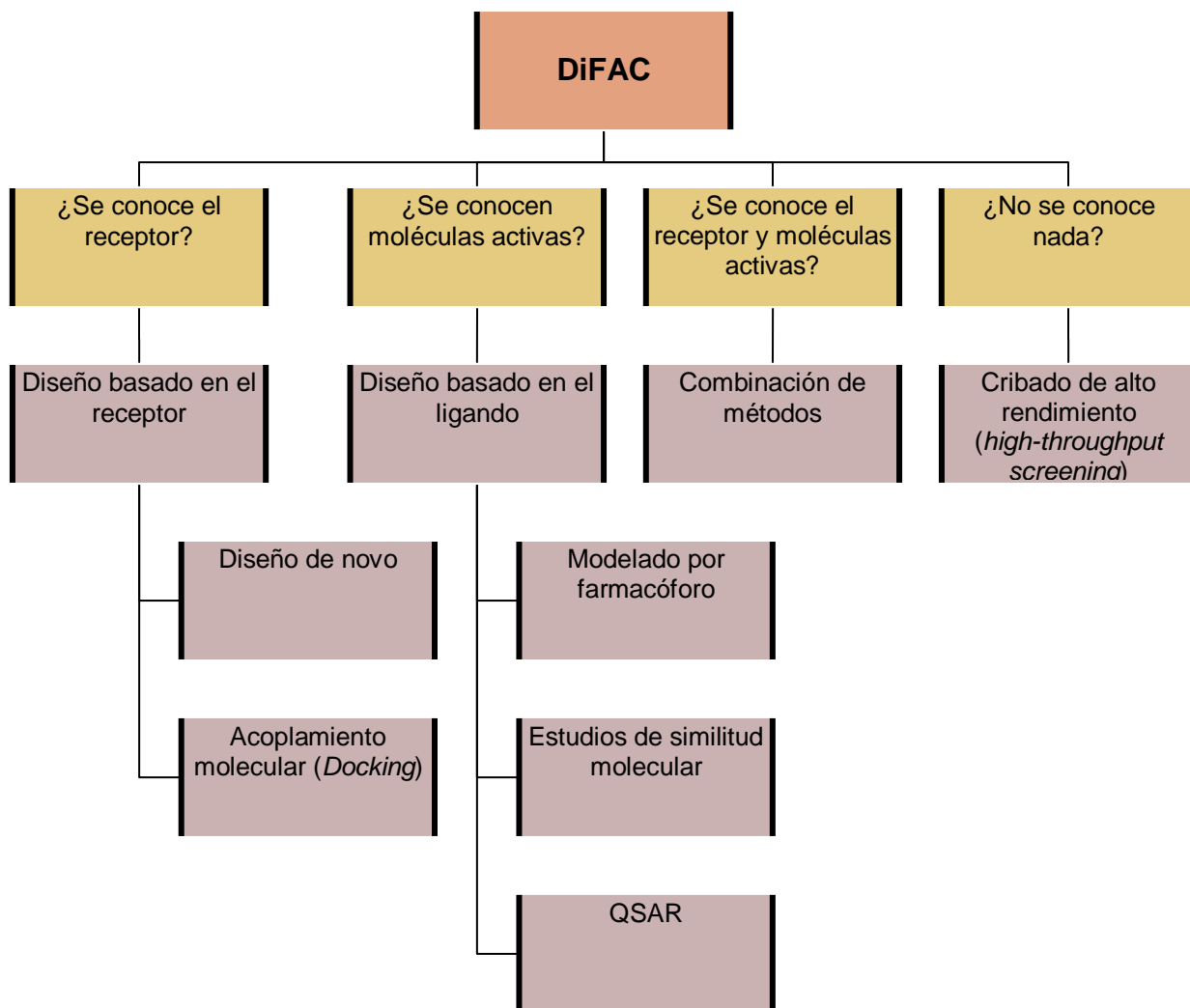


Figura 11. Clasificación de algunas metodologías usadas en DiFAC.

2.3.2. Estudios QSAR-3D previos sobre bencimidazoles con actividad antiprotozoaria

Durante los años noventa comenzaron a adquirir popularidad dos metodologías de QSAR-3D llamadas Análisis Comparativo de Campos Moleculares (CoMFA) [41] y Análisis Comparativo de Índices de Similitud (CoMSIA) [42–44]. Estos métodos se caracterizan por ofrecer un entorno gráfico que facilita el estudio de las relaciones estructura-actividad; además, un gran número de trabajos sustentan su amplio potencial predictivo. Por ejemplo, en diversos trabajos se aplicaron estos métodos para el estudio de derivados de bencimidazol con actividad antitumoral, antiviral y agonista 5HT₄ [45–47]. Por otra parte, en nuestro grupo de investigación, recientemente, se estudiaron derivados de bencimidazol con actividad contra los

protozoarios *E. histolytica* y *G. intestinalis*. En particular, contra *E. histolytica* se obtuvieron modelos CoMFA con alto poder predictivo [48], esto es, cumplen con las condiciones $r^2 > 0.6$ (coeficiente de correlación para el grupo de trabajo), $q^2 > 0.5$ (correlación determinada por el método *leave-one-out* para validación interna), $R^2 > 0.6$ (coeficiente de correlación para el grupo de prueba) y otros criterios estrictos de validación [30, 49–50]. El mejor modelo obtenido mostró buena correlación ($r^2 = 0.873$) y una excelente capacidad predictiva ($q^2 = 0.720$ y $R^2 = 0.951$). De este modelo se obtuvieron los mapas de contorno que se muestran en la Figura 12. La principal conclusión de ese trabajo fue que el aumento en el tamaño del sustituyente en posición 5 incrementa la actividad amebicida, mientras que las cargas parciales positivas en posición 2 también favorecen la actividad [48].



Figura 12. Mapas de contornos para el mejor modelo CoMFA construido con derivados de bencimidazol activos contra *E. histolytica*. Las regiones donde se incrementa la actividad con sustituyentes voluminosos están en verde y donde se disminuye en amarillo. Las regiones donde la actividad se favorece por cargas atómicas parciales negativas están en rojo y donde se desfavorece en azul.

También se construyó un modelo CoMFA de derivados de bencimidazol con actividad contra *G. intestinalis* que mostró valores modestos de correlación y capacidad predictiva ($r^2 = 0.588$ y $q^2 = 0.487$, respectivamente) [51]; sin embargo, el modelo proporcionó información importante en el mapa de contornos presentado en la Figura 13. Las principales conclusiones de ese trabajo fueron que los sustituyentes de mayor tamaño en posición 2 aumentan la actividad, mientras que en posición 5 tienen un efecto desfavorable; además se observó el efecto desfavorable sobre la actividad causado por cargas parciales negativas (o favorable causado por cargas atómicas parciales positivas) sobre sustituyentes en las posiciones 5 y 6 [51].

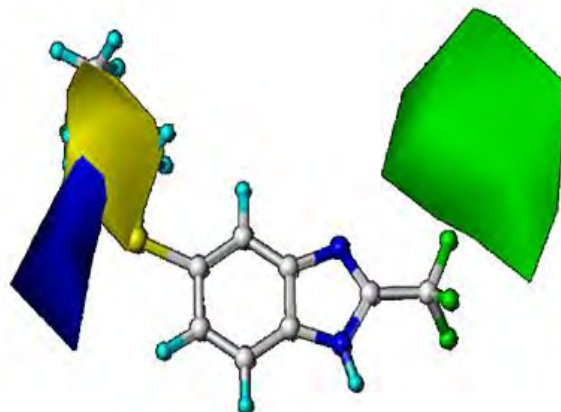


Figura 13. Mapa de contornos para el mejor modelo CoMFA construido con derivados de bencimidazol activos contra *G. intestinalis*. El código de colores tiene el mismo significado que en la Figura 12.

Hasta ese momento se contaba con estudios de relaciones estructura-actividad contra *E. histolytica* y *G. intestinalis*. Aunque también se contaba con una cantidad importante de información sobre la actividad tricomonocida de derivados de bencimidazol, no se habían hecho estudios computacionales; por lo que en esta tesis se estudiaron las relaciones estructura-actividad contra *T. vaginalis* usando las metodologías de CoMFA y CoMSIA.

2.3.3. Panorama de actividad y su utilidad en la construcción de modelos QSAR

Basados en el principio de similitud, se asume que moléculas similares tendrán actividades similares; sin embargo, esto no siempre se cumple y es común encontrar moléculas similares que tienen actividades muy diferentes, aunque este comportamiento contradice el principio de similitud, es un hecho real que llevó a introducir el concepto de *activity cliff*. Los *activity cliffs* se definen por la relación entre la diferencia de actividad para dos compuestos y su distancia en un espacio químico dado [52]. Los *activity cliffs* han sido el tema central de numerosos trabajos [52–54], además se les ha considerado como puntos clave en la optimización de moléculas activas. Recientemente se introdujo el concepto de panorama de actividad (*activity landscape*) donde puede representarse de manera conjunta el espacio químico y la actividad biológica. Una de las características de las representaciones del panorama de actividad es que facilitan la detección de los *activity cliffs*. La utilidad de detectar *activity cliffs* se basa en que representan puntos fuera de correlación en estudios

QSAR. Sin embargo, los *activity cliffs* brindan información valiosa para optimizar la actividad biológica, ya que con pequeños cambios estructurales se logran cambios grandes en la actividad biológica [52].

Uno de los métodos usados para el estudio del panorama de actividad emplea los mapas SAS (structure-activity similarity) para obtener una caracterización visual y cuantitativa del perfil estructura-actividad [53]. La forma general de los mapas SAS se muestra en la Figura 14. Los mapas SAS se construyen graficando, para cada par de compuestos, la similitud en estructura en el eje X contra la similitud en actividad en el eje Y [55].

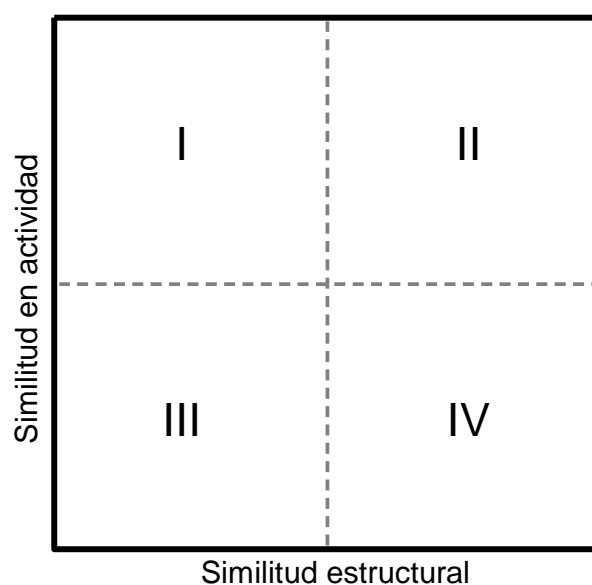


Figura 14. Forma esquemática un mapa SAS dividido en cuatro regiones (I–IV). La región I muestra diferencias en la estructura que dirigen a diferencias pequeñas en actividad; la región II representa el SAR continuo, la región III brinda poca información y la región IV representa el SAR discontinuo (*activity cliffs*).

En estos mapas se pueden distinguir cuatro regiones (I–IV). Los puntos que caen sobre la región I corresponden a pares de moléculas con alta similitud en actividad y baja similitud estructural y por lo tanto pueden ser asociados con cambios del núcleo base (*scaffold hopping*) o cambios grandes en cadenas laterales. Los puntos en la región II se refieren a moléculas con alta similitud estructural y en actividad, y pueden ser asociados con un comportamiento continuo en las relaciones estructura-actividad (cumplen el principio de similitud). La región III corresponde a pares de moléculas con baja similitud estructural y en actividad; esta región se

considera poco informativa. La región IV contiene pares de moléculas con alta similitud estructural y baja similitud en actividad y se asocian con relaciones estructura-actividad discontinuas y la presencia de “*activity cliffs*”.

3. JUSTIFICACIÓN

Considerando la problemática de las enfermedades parasitarias producidas por protozoarios, aunado a la carencia de fármacos seguros, es importante realizar investigación básica con el propósito de contribuir en el desarrollo de nuevas moléculas activas. En este trabajo se emplea la base de datos de bencimidazoles existente y algunas herramientas computacionales con el fin de estudiar las características estructurales necesarias para optimizar la actividad antiprotozoaria y finalmente diseñar y sintetizar un grupo de nuevos compuestos con potente actividad antiprotozoaria.

4. HIPÓTESIS

Los modelos computacionales, generados a partir de los derivados de bencimidazol, mostrarán información sobre los requerimientos estructurales en este tipo de compuestos para la actividad antiprotozoaria que será de gran utilidad para encontrar nuevos compuestos activos.

5. OBJETIVOS

5.1. *Objetivo general*

Sintetizar un grupo de derivados de bencimidazol con actividad antiprotozoaria diseñados mediante estudios de relaciones estructura-actividad.

Para cumplir con el objetivo general se realizaron las siguientes actividades:

- Analizar las relaciones estructura-actividad de derivados de bencimidazol con actividad tricomonida y giardicida aplicando técnicas de quimioinformática.
- Construir y validar una serie de modelos QSAR-3D usando la base de datos de bencimidazoles tricomonidas.
- Diseñar una serie de moléculas empleando las conclusiones obtenidas de los estudios SAR y QSAR.

JUSTIFICACIÓN, HIPÓTESIS Y OBJETIVOS

- De las moléculas diseñadas, sintetizar un grupo representativo.
- Caracterizar las moléculas sintetizadas empleando métodos espectroscópicos y espectrométricos.
- Evaluar la actividad in vitro sobre *T. vaginalis*, *G. intesinalis* y *E. histolytica*.
- Comparar los datos de actividad tricomonocida calculada y experimental de las moléculas sintetizadas.

6. METODOLOGÍA

El presente trabajo está organizado en cuatro secciones. La primera abarca estudios sobre el panorama de actividad tricomonocida y giardicida de derivados de bencimidazol; en la segunda se aborda el estudio de las relaciones cuantitativas estructura-actividad de derivados de bencimidazol con actividad tricomonocida; la tercera sección trata sobre la actividad antiprotozoaria de los inhibidores de la bomba de protones y, por último, en la cuarta sección se describe el diseño y la síntesis de nuevos derivados de bencimidazol pronosticados con buena actividad antiprotozoaria. Es muy importante mencionar que, por simplicidad, la numeración de los compuestos estudiados o sintetizados es propia de cada sección.

6.1. Panorama de actividad tricomonocida y giardicida de derivados de bencimidazol

6.1.1. Base de datos

Las estructuras y datos de actividad biológica para 55 derivados de bencimidazol considerados en este estudio se tomaron de reportes recientes de nuestro grupo de investigación y otro grupo externo (Tabla 2) [16–19, 21]. Estas moléculas poseen datos de actividad contra los protozoarios *T. vaginalis* y *G. intestinalis* determinadas en el mismo laboratorio bajo las mismas condiciones. Para *T. vaginalis*, las actividades van desde 2 nM ($pCl_{50} = 8.70$) hasta 29,512 nM ($pCl_{50} = 4.53$). Para *G. intestinalis* las actividades van desde 20.4 nM ($pCl_{50} = 7.69$) hasta 10,471 nM ($pCl_{50} = 4.98$).

6.1.2. Búsqueda conformacional

Los cálculos de modelado molecular se realizaron usando el programa Spartan'02 [56], sobre el sistema operativo Red Hat Enterprise Linux 4. El análisis conformacional se realizó mediante el método de búsqueda sistemático usando torsiones de 60° por cada enlace rotatable y el campo de fuerza MMFF94. La búsqueda se restringió a las mejores conformaciones en un rango de 10 kcal/mol relativo a la conformación de menor energía. Los confórmeros obtenidos y las moléculas sin

enlaces rotables fueron posteriormente optimizadas usando el método semiempírico PM3 y se eligió el confórmero de mínima energía [57].

6.1.3. Similitud estructural 2D y 3D

Para cada molécula se calcularon las representaciones moleculares implementadas en programa MOE [58]: MACCS keys 166 bits, GpiDAPH3, TGD y TGT. También se usaron las representaciones implementadas en programa Canvas [31, 59]: Radial, dendritic, atom pairs y MOLPRINT 2D. Además se calcularon representaciones en tres dimensiones considerando la conformación de mínima energía calculada mediante el método descrito en la sección anterior (6.1.2). Se calcularon representaciones basadas en 3 y 4 puntos de farmacóforo: piDAPH3 y piDAPH4, implementadas en MOE; y 3- y 4-point pharmacophores, implementadas en Canvas.

Para las moléculas m_i y m_j se calculó la similitud estructural por pares usando el coeficiente de Tanimoto mediante la siguiente ecuación [60–61]:

$$T_{ij} = \frac{C_{ij}}{C_i + C_j - C_{ij}} \quad (1)$$

donde T_{ij} es la similitud estructural entre la i -ésima y la j -ésima molécula; C_{ij} es el número total de características en común; C_i es el número total de características en la i -ésima molécula y C_j es el número total de características en la j -ésima molécula.

6.1.4. Similitud usando propiedades calculadas

Las propiedades usadas en este estudio se calcularon usando el programa Canvas e incluyen: peso molecular, número de enlaces rotables, aceptores de hidrógeno, donadores de hidrógeno, superficie polar y coeficiente de partición octanol/agua. Las propiedades se escalaron usando la siguiente ecuación:

$$p_{ki} = \frac{P_{ki} - \overline{P}_k}{\sigma_{P_k}} \quad (2)$$

donde p_{ki} es la versión escalada de la k -ésima propiedad para la i -ésima molécula, P_{ki} representa la propiedad no escalada, y \overline{P}_k y σ_{P_k} denotan el promedio y la desviación

estándar, respectivamente, para la k -ésima propiedad sobre todas las moléculas en el estudio.

La distancia euclidiana entre las propiedades para cada par de moléculas se calculó mediante la siguiente ecuación [61]:

$$d_{ij} = \left[\sum_{k=1}^K (p_{ki} - p_{kj})^2 \right]^{1/2} \quad (3)$$

donde d_{ij} es la distancia euclidiana entre las propiedades para la i -ésima y la j -ésima moléculas, y p_{ki} y p_{kj} son el valor de la propiedad escalada k para la i -ésima y j -ésima moléculas, respectivamente. Para este estudio el valor de k es 6 (seis propiedades).

Posteriormente, las distancias euclidianas se escalaron de 0 a 1 usando la siguiente ecuación:

$$sd_{ij} = \frac{d_{ij} - \min d_{ij}}{\max d_{ij} - \min d_{ij}} \quad (4)$$

donde sd_{ij} es la distancia escalada, y $\max d_{ij}$ y $\min d_{ij}$ indican el rango de distancias en la base de datos.

La similitud en propiedades para cada par de moléculas se calculó con la ecuación:

$$PS_{ij} = 1 - sd_{ij} \quad (5)$$

donde PS_{ij} es la similitud en propiedades para la i -ésima y la j -ésima moléculas, y sd_{ij} es la distancia escalada.

6.1.5. Similitud en actividad

Para cada par de moléculas la similitud en actividad para *T. vaginalis* y *G. intestinalis* se midió usando la siguiente ecuación:

$$AS_{ij} = 1 - \frac{|A_i - A_j|}{\max - \min} \quad (6)$$

donde A_i y A_j son las actividades en pCl_{50} para la i -ésima y j -ésima moléculas y \max - \min indican el rango entre el valor máximo y mínimo de actividad en la base de datos.

6.1.6. Mapas SAS

Para cada protozario se construyeron los mapas SAS graficando, para cada par de compuestos, la similitud estructural en el eje X contra la similitud en actividad en el eje Y [55, 62]. Para caracterizar los mapas SAS obtenidos con diferentes medidas de similitud, cada mapa se dividió en cuatro regiones (I–IV). Se usaron límites de 0.5 y 0.75 de similitud en actividad (eje Y) para separar las regiones I y II de III y IV, respectivamente. Para establecer el límite sobre similitud estructural (eje X) se usó la mediana de la similitud estructural de los compuestos más activos.

Para comparar entre los mapas SAS generados por medio de diferentes representaciones de similitud estructural se calculó el grado de consenso (*DoC*) [55]. *DoC* mide el número de datos que coinciden en una misma región para representaciones diferentes y se calcula con la siguiente ecuación:

$$DoC_{m,n}^R = \frac{Cp_{m,n}}{p_m + p_n - Cp_{m,n}} \quad (7)$$

donde $Cp_{m,n}$ es el número de pares consenso en la región R ($R = I-IV$) entre los métodos m y n ; p_m es el número de pares de moléculas asignado por el método m en la región R y p_n es el número de pares de moléculas asignado por el método n en la misma región. Es importante notar que el *DoC* depende de los límites establecidos para definir las cuatro regiones.

6.1.7. Mapas SAS consenso

Para generar los mapas SAS consenso se combinaron diferentes representaciones de similitud estructural. En este trabajo se consideró el promedio de cuatro representaciones no correlacionadas (radial, MACCS, TGD y piDAPH3), no obstante, pueden usarse otras representaciones. Los mapas SAS consenso se construyeron graficando la similitud en actividad contra el promedio de los valores de similitud de las cuatro representaciones seleccionadas.

6.1.8. Mapas DAD

En este trabajo se proponen por primera vez los mapas de diferencia de actividad dual (DAD), los cuales ayudan a explorar de una manera sistemática las diferencias de actividad frente a dos parásitos. Los mapas DAD se construyeron graficando las diferencias pareadas de actividad para *T. vaginalis* contra las diferencias pareadas de actividad para *G. intestinalis* calculadas mediante la siguiente ecuación:

$$\Delta pCl_{50}(P)_{ij} = pCl_{50}(P)_i - pCl_{50}(P)_j \quad (8)$$

donde $pCl_{50}(P)_i$ y $pCl_{50}(P)_j$ son las actividades para la *i*-ésima y *j*-ésima moléculas contra el parásito P. Para las comparaciones duplicadas, por ejemplo el par **1_2** y **2_1** que tienen el mismo valor pero diferente signo, sólo se consideró una de ellas, ya que proporcionan la misma información. Finalmente a estos mapas se les agregó información estructural por medio de representaciones de similitud.

6.2. Relaciones cuantitativas estructura-actividad de derivados de bencimidazol con actividad tricomonocida

6.2.1. Base de datos y actividad biológica

La base de datos usada en este estudio contiene 70 derivados de bencimidazol con actividad tricomonocida previamente publicados (Tabla 12). Todos los compuestos han sido sintetizados en nuestro grupo de investigación [11, 16–17, 20–21], excepto **24–45**, reportados por Andrzejewska et al. [18–19]. Cabe mencionar que todos los ensayos biológicos se realizaron en el mismo laboratorio usando la cepa GT3 de *Trichomonas vaginalis*. Los valores de Cl_{50} se convirtieron a $pCl_{50} = -\log Cl_{50}$, usando la actividad en unidades de mol/L.

6.2.2. Selección del grupo de prueba y de trabajo

El grupo inicial de 70 compuestos se dividió en un grupo de trabajo con 57 moléculas (81%) y un grupo de prueba con 13 moléculas (19%). El rango de actividad para el grupo de trabajo abarca desde 2 nM ($pCl_{50} = 8.70$) hasta 29,512 nM ($pCl_{50} =$

4.53); mediana $pCl_{50} = 6.70$. El rango de actividad para el grupo de prueba abarca desde 20 nM ($pCl_{50} = 7.69$) hasta 28,440 nM ($pCl_{50} = 4.55$); mediana $pCl_{50} = 6.98$. Es importante mencionar que los compuestos para ambos grupos se seleccionaron de tal manera que existiera diversidad estructural y un rango de actividad biológica similar [30, 49–50].

6.2.3. Preparación de estructuras

La búsqueda conformacional y optimización de geometría se realizó mediante la metodología previamente descrita (sección 6.1.2). Como siguiente paso se calcularon las cargas atómicas (Mulliken y electrostáticas) usando el método semiempírico PM3 implementado en Spartan'02 [56, 63–64]. Las cargas atómicas de Gasteiger-Hückel [65] se calcularon usando el programa Sybyl 7.3 [66].

Para la selección de conformeros se usaron dos metodologías. Para la primera aproximación se seleccionaron los conformeros de mínima energía obtenidos empleando el método semiempírico PM3. Como método alternativo se seleccionaron conformeros usando similitud 3D. Para este último caso se usó el método de superposición rápida de estructuras químicas implementado en el programa ROCS [67]. La superposición de estructuras químicas se llevó a cabo usando como referencia a los compuestos **9**, **10** y **45** en su conformación de mínima energía. Para seleccionar los conformeros en base a su mayor valor de similitud respecto a su molécula de referencia se usó el puntaje combinado (combo score). Este puntaje está compuesto por el índice de Tanimoto basado en traslape de volumen y el puntaje de color (color score). El índice de Tanimoto se calculó mediante la ecuación (1) considerando que C_i es el volumen de la i -ésima molécula, C_j volumen de la j -ésima molécula y C_{ij} es el volumen de traslape entre las i -ésima y la j -ésima moléculas [61, 68]. El puntaje de color se calculó con un campo de fuerza que incluye parámetros para donadores y aceptores de hidrógeno, grupos hidrófobos, aniones, cationes y anillos [68–69].

6.2.4. Alineamiento molecular

Todos los compuestos fueron alineados usando el núcleo de bencimidazol como referencia, el cual es la subestructura común en todas las moléculas. Dado que

el mecanismo de acción y blanco biológico de este tipo de moléculas es aún desconocido, se asume que todos los compuestos tienen el mismo mecanismo de acción y modo de unión.

6.2.5. Selección de formas tautoméricas

Dado que muchas de las moléculas en el presente estudio contienen dos o tres formas tautoméricas importantes, se seleccionaron formas tautoméricas representativas para generar los modelos usando valores estimados de la actividad mediante cálculos QSAR-3D [48]. Primero se construyó un modelo que contiene moléculas sin formas tautoméricas; posteriormente, este modelo se usó para predecir la actividad de aquellas moléculas con tautómeros. Se eligieron los tautómeros cuya actividad calculada fue la más cercana a la actividad experimental para usarse en los modelos QSAR generados posteriormente.

6.2.6. Metodología CoMFA

Los estudios CoMFA se realizaron empleando el programa Sybyl 7.3. Para el cálculo de los descriptores estéricos y electrostáticos se usó una rejilla cúbica de 2 Å con espacio equidistante entre sus puntos y se extendió por 4 Å más allá de la posición ocupada por las moléculas alineadas en todas las direcciones. Los valores de los descriptores estéricos y electrostáticos se calcularon, en cada punto de la rejilla, usando un átomo de carbono sp^3 con carga +1, radio de 1 Å y una constante dieléctrica dependiente de la distancia. Los cálculos para ambos parámetros se truncaron a un valor de energía de 30 kcal/mol [30, 32–33].

Los descriptores estéricos se calcularon usando el potencial de Lennard-Jones mediante la ecuación [30, 32–33]:

$$E_{vdW} = \sum_{i=1}^n \frac{A_i}{r_{ij}^{12}} - \frac{B_j}{r_{ij}^6} \quad (9)$$

donde E_{vdW} es la energía de interacción de Van der Waals; A y B son constantes que dependen del radio de Van der Waals de los átomos correspondientes; r_{ij} es la distancia entre el átomo i y el átomo de prueba j .

Los descriptores electrostáticos se calcularon mediante el potencial de Coulomb usando la siguiente ecuación [30, 32–33]:

$$E_C = \sum_{i=1}^n \frac{q_i q_j}{D r_{ij}} \quad (10)$$

donde E_C es la energía de interacción de Coulomb; q_i es la carga parcial para el átomo i , q_j es la carga parcial para el átomo de prueba j ; D es la constante dieléctrica; r_{ij} es la distancia entre el átomo i y el átomo de prueba j .

6.2.7. Metodología CoMSIA

Se usaron índices de similitud basados en cinco diferentes propiedades [42–44]. Los descriptores estéricos, electrostáticos, hidrófobos, donadores y aceptores de puentes de hidrógeno se generaron usando la misma rejilla definida para los modelos de CoMFA y empleando el átomo de carbono sp^3 con carga +1 y radio de 1 Å. Los índices de similitud ($A_{F,k}$) para la molécula j y el átomo i en el punto de la rejilla q se calcularon usando la siguiente ecuación:

$$A_{F,k}^q(j) = -\sum_{i=1}^n \omega_{prueba,k} \omega_{ik} e^{-\alpha r_{iq}^2} \quad (11)$$

donde ω_{ik} es el valor real de la propiedad fisicoquímica k del átomo i ; $\omega_{prueba,k}$ es el átomo de prueba con carga +1 y radio de 1 Å, carácter hidrófobo +1, propiedad de donador y aceptor de hidrógeno +1; $-\alpha$ es el factor de atenuación; r_{iq} es la distancia mutua entre el átomo de prueba en el punto q de la rejilla y el átomo i de la molécula a evaluar. Los índices estéricos están relacionados a la tercera potencia del radio atómico; los descriptores electrostáticos se derivaron de las cargas atómicas [42]; los descriptores hidrófobos de parámetros basados en átomos [44]; mientras que los índices de donadores y aceptores de puentes de hidrógeno se obtuvieron del método basado en reglas generadas por resultados experimentales [43]. Se usó el valor preestablecido de 0.3 como factor de atenuación.

6.2.8. Análisis de mínimos cuadrados parciales

Se empleó el análisis de mínimos cuadrados parciales (PLS) implementado en Sybyl 7.3 para generar una regresión que correlaciona los descriptores con la actividad

tricomonicida. Se usó la validación cruzada para obtener el número óptimo de componentes principales mediante la metodología *Leave-One-Out* (LOO). El coeficiente de correlación q^2 se calculó mediante la siguiente ecuación:

$$q^2 = 1 - \frac{\sum (Y_{pred} - Y_{exp})^2}{\sum (Y_{exp} - Y_{prom})^2} \quad (12)$$

donde Y_{pred} , Y_{exp} y Y_{prom} son los valores para la actividad calculada, experimental y promedio, respectivamente.

6.2.9. Validación

Para la validación de los modelos se consideró el coeficiente de correlación de validación interna q^2 y las predicciones para el grupo de prueba de 13 compuestos. Se calcularon las actividades tricomonicidas del grupo de prueba usando los modelos generados y, por análisis de regresión, se determinaron los valores necesarios para aplicar las condiciones propuestas por Golbraikh y Tropsha [49–50]:

- I. $q^2 > 0.5$
- II. $R^2 > 0.6$
- III. $[(R^2 - R_0^2)/R^2] < 0.1$ ó $[(R^2 - R_0'^2)/R^2] < 0.1$
- IV. $0.85 \leq k \leq 1.15$ ó $0.85 \leq k' \leq 1.15$

donde, q^2 es el coeficiente de correlación para la validación cruzada LOO; R^2 es el coeficiente de correlación para las actividades experimentales (y) contra actividades predichas (\tilde{y}) para el grupo de prueba; R_0^2 y $R_0'^2$ son los coeficientes de correlación para la línea de regresión que pasa por el origen para y contra \tilde{y} y \tilde{y} contra y respectivamente; k y k' son las pendientes para las líneas de regresión que pasan por el origen $y^{r0} = k\tilde{y}$ y $\tilde{y}^{r0} = k'y$, las cuales se calculan mediante las siguientes ecuaciones:

$$k = \frac{\sum y_i \tilde{y}_i}{\tilde{y}_i^2} \quad (13)$$

$$k' = \frac{\sum y_i \tilde{y}_i}{y_i^2} \quad (14)$$

Los modelos que satisfacen las cuatro condiciones mencionadas anteriormente (I–IV) se consideran predictivos y cercanos a ser modelos ideales.

6.3. Actividad antiprotozoaria de los inhibidores de la bomba de protones

6.3.1. Modelado molecular

Las estructuras de los inhibidores de la bomba de protones y de las moléculas de referencia se sometieron a un análisis conformacional en el programa Spartan'02 usando el método de Montecarlo con incrementos de 30° por cada enlace rotatable y el campo de fuerza MMFF94. La geometría de los conformeros obtenidos en la búsqueda conformacional se optimizó usando el método Hartree-Fock con funciones base 6-31G* y a las estructuras resultantes se les calculó el potencial electrostático. La alineación de las moléculas y medición de distancias respecto a posibles puntos farmacofóricos se realizó en el programa Sybyl 7.3.

6.3.2. Obtención de los compuestos

Los compuestos se obtuvieron de sus presentaciones comerciales por extracción. El omeprazol se obtuvo de la presentación comercial "Losec ATM" producida por AstraZeneca AB (Suecia), el lansoprazol se obtuvo de "SeraxTM" producido por Landsteiner Scientific S.A. de C.V. (México), el rabeprazol se obtuvo de "ParietTM" producido por Eisai Co. Ltd. (Japon) y el pantoprazol de "PantozolTM" producido por Nycomedn Oranienburg GmbH (Alemania). Los compuestos obtenidos se analizaron mediante cromatografía en capa fina y se les determinó su punto de fusión; además se caracterizó su estructura usando espectroscopia de resonancia magnética nuclear y espectrometría de masas.

6.3.3. Evaluaciones biológicas

En este trabajo se evaluó la actividad antiprotozoaria (*in vitro*) de los inhibidores de la bomba de protones contra *G. intestinalis*, *E. histolytica* y *T. vaginalis*.

Para estos ensayos se siguió el método descrito por Cedillo-Rivera et al. y Chávez et al. [8, 10]. El procedimiento general aplicado a cada parásito varió solamente en el número de trofozoitos incubados correspondiendo a 6×10^3 para *E. histolytica* y 5×10^4 para *G. intestinalis* y *T. vaginalis*.

Se incubó el número de trofozoitos correspondiente, según el tipo de parásito a evaluar, con los diferentes compuestos a concentraciones de 0.005, 0.01, 0.05, 0.1, 0.5 y 1 µg/mL. El albendazol y metronidazol se usaron como controles positivos y el DMSO libre de fármacos como control negativo. Los tubos se incubaron a 37 °C por 48 h. Posteriormente, los trofozoitos se lavaron y resembraron en medio fresco por 48 h más. Finalmente, se determinó el número de trofozoitos en una cámara de Neubauer. Los resultados se analizaron por medio de un análisis probit para determinar la concentración inhibitoria 50 (CI₅₀), de cada uno de los compuestos, que corresponde a la concentración del compuesto que inhibió en 50% el crecimiento de los trofozoitos.

6.4. Diseño, síntesis y actividad antiprotozoaria de nuevos derivados de bencimidazol

6.4.1. Moléculas diseñadas

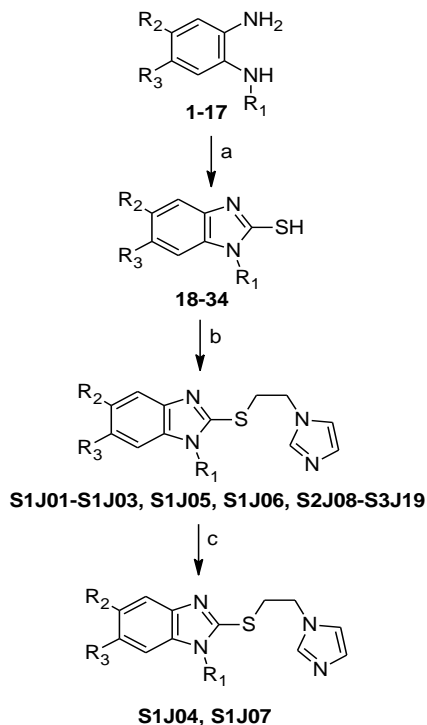
En base a los estudios anteriores de relaciones estructura-actividad se diseñó una serie de derivados de bencimidazol y se eligieron 19 para su síntesis y evaluación antiprotozoaria (Tabla 1).

Tabla 1. Moléculas diseñadas

Clave	R ₁	R ₂	R ₃
S1J01	H	H	H
S1J02	H	Cl	H
S1J03	H	Cl	Cl
S1J04	CH ₃	H	H
S1J05	CH ₃	H	Cl
S1J06	CH ₃	Cl	H
S1J07	CH ₃	Cl	Cl
S2J08	H	COOCH ₃	H
S2J09	H	COOCH ₃	Cl
S2J10	CH ₃	H	COOCH ₃
S2J11	CH ₃	Cl	COOCH ₃
S2J12	CH ₃	COOCH ₃	H
S2J13	CH ₃	COOCH ₃	Cl
S3J14	H	OCH ₂ CH ₃	H
S3J15	H	OCH ₂ CH ₃	Cl
S3J16	CH ₃	H	OCH ₂ CH ₃
S3J17	CH ₃	Cl	OCH ₂ CH ₃
S3J18	CH ₃	OCH ₂ CH ₃	H
S3J19	CH ₃	OCH ₂ CH ₃	Cl

6.4.2. Síntesis de los compuestos diseñados

Los compuestos diseñados en este trabajo se sintetizaron a partir del 1,2-diaminobenceno adecuadamente sustituido (**1–17**) (Esquema 1).



- 1, 18, S1J01:** $\text{R}_1=\text{R}_2=\text{R}_3=\text{H}$
2, 19, S1J02: $\text{R}_1=\text{H}, \text{R}_2=\text{Cl}, \text{R}_3=\text{H}$
3, 20, S1J03: $\text{R}_1=\text{H}, \text{R}_2=\text{R}_3=\text{Cl}$
S1J04: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{R}_3=\text{H}$
4, 21, S1J05: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{H}, \text{R}_3=\text{Cl}$
5, 22, S1J06: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{Cl}, \text{R}_3=\text{H}$
S1J07: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{R}_3=\text{Cl}$
6, 23, S2J08: $\text{R}_1=\text{H}, \text{R}_2=\text{COOCH}_3, \text{R}_3=\text{H}$
7, 24, S2J09: $\text{R}_1=\text{H}, \text{R}_2=\text{COOCH}_3, \text{R}_3=\text{Cl}$
8, 25, S2J10: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{H}, \text{R}_3=\text{COOCH}_3$
9, 26, S2J11: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{Cl}, \text{R}_3=\text{COOCH}_3$
10, 27, S2J12: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{COOCH}_3, \text{R}_3=\text{H}$
11, 28, S2J13: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{COOCH}_3, \text{R}_3=\text{Cl}$
12, 29, S3J14: $\text{R}_1=\text{H}, \text{R}_2=\text{OCH}_2\text{CH}_3, \text{R}_3=\text{H}$
13, 30, S3J15: $\text{R}_1=\text{H}, \text{R}_2=\text{OCH}_2\text{CH}_3, \text{R}_3=\text{Cl}$
14, 31, S3J16: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{H}, \text{R}_3=\text{OCH}_2\text{CH}_3$
15, 32, S3J17: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{Cl}, \text{R}_3=\text{OCH}_2\text{CH}_3$
16, 33, S3J18: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{OCH}_2\text{CH}_3, \text{R}_3=\text{H}$
17, 34, S3J19: $\text{R}_1=\text{CH}_3, \text{R}_2=\text{OCH}_2\text{CH}_3, \text{R}_3=\text{Cl}$

(a) CS_2 , KOH, EtOH- H_2O , 40-50 °C, 8 h; (b) 1-(2-cloroetil)-1H-imidazol, CH_3CN , KOH, o DBU, 60 °C, 8 h; (c) CH_3I , acetona, K_2CO_3 .

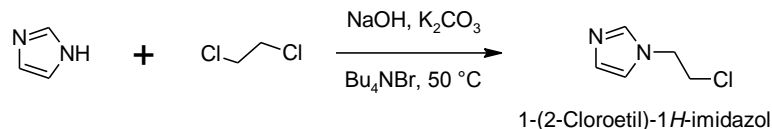
Esquema 1. Reactivos y condiciones para la síntesis de los compuestos **S1J01–S3J19**.

El primer paso consistió en la ciclación del 1,2-diaminobenceno apropiado con disulfuro de carbono para formar los 2-mercaptobencimidazoles **18–34** [70]. Posteriormente, los 2-mercaptobencimidazoles se trataron con 1-(2-cloroetil)-1H-imidazol para generar los compuestos de las series **S1**, **S2** y **S3**; con excepción de los compuestos **S1J04** y **S1J07** que se obtuvieron por metilación de **S1J01** y **S1J03**, respectivamente.

El intermediario común 1-(2-cloroetil)-1H-imidazol y los 1,2-diaminobencenos **2–17** no estaban disponibles comercialmente, por lo que se prepararon a partir de compuestos comerciales. La síntesis de éstos se describe en las siguientes secciones.

6.4.2.1. 1-(2-Cloroetil)-1H-imidazol

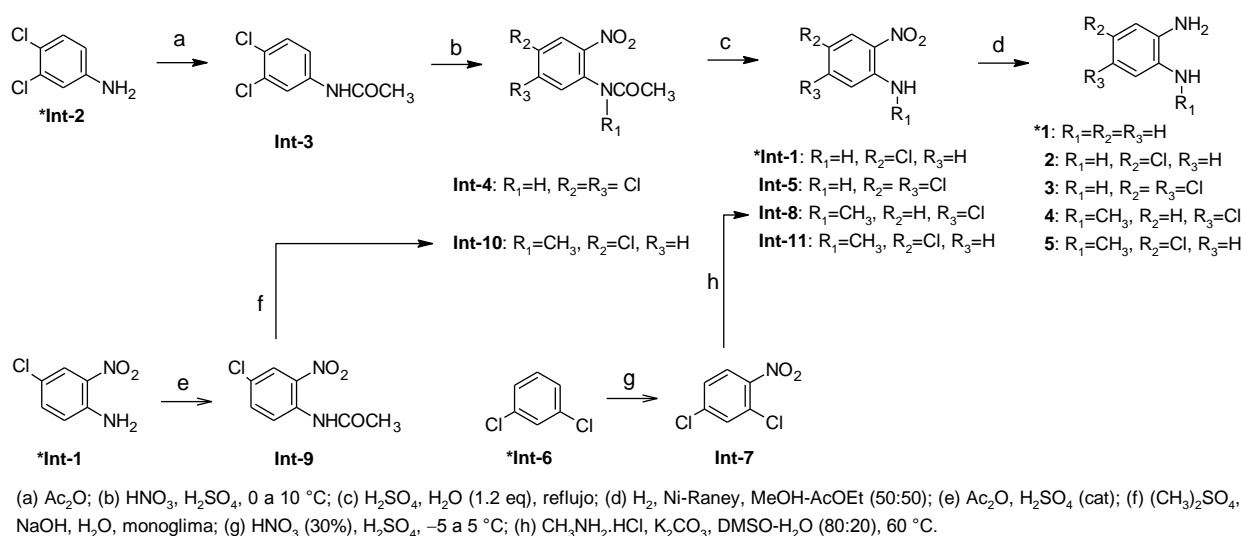
El intermediario 1-(2-cloroetil)-1H-imidazol se sintetizó empleando una reacción de sustitución nucleofílica entre el imidazol y el 1,2-dicloroetano (Esquema 2) [71].



Esquema 2. Preparación del 1-(2-cloroetil)-1H-imidazol.

6.4.2.2. 1,2-Diaminobencenos 1–5 (Serie 1)

La preparación de los 1,2-diaminobencenos que se utilizaron para la síntesis de los compuestos de la Serie 1 (**S1J01–S1J07**) se realizó partiendo de materias primas comerciales (Esquema 3). El 1,2-diaminobenceno (*o*-fenilendiamina) **1** se encontró disponible comercialmente, mientras que los 1,2 diaminobencenos **2–5** se sintetizaron mediante la ruta que se muestra en el Esquema 3. Las reacciones utilizadas incluyen procesos de nitración, sustitución nucleofílica aromática, acetilación, metilación, hidrólisis e hidrogenación catalítica.

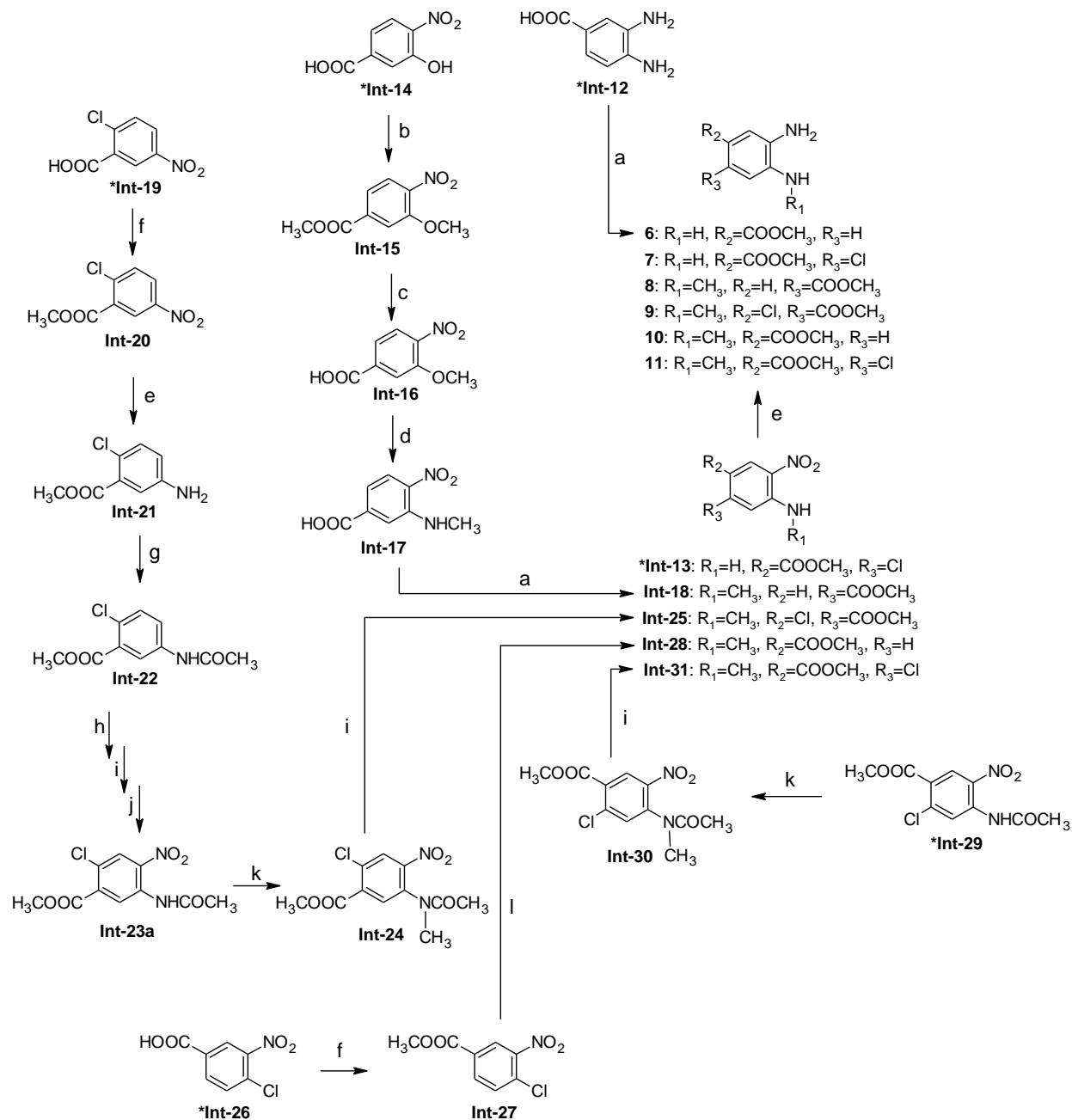


Esquema 3. Síntesis de los 1,2-diaminobencenos 2–5 a partir de compuestos comerciales (*).

6.4.2.3. 1,2-Diaminobencenos 6–11 (Serie 2)

La síntesis de los 1,2-diaminobencenos **6–11**, usados para la preparación de los compuestos de la Serie 2 (**S2J08–S2J13**), se realizó a partir de los ácidos benzoicos o benzoatos de metilo disponibles comercialmente. La síntesis general se muestra en el Esquema 4. Cabe mencionar que el intermediario **Int-23a** se obtuvo a partir de **Int-22**

por una serie de tres pasos que facilitaron su purificación, los detalles se discutirán en la sección de resultados y discusión correspondiente.

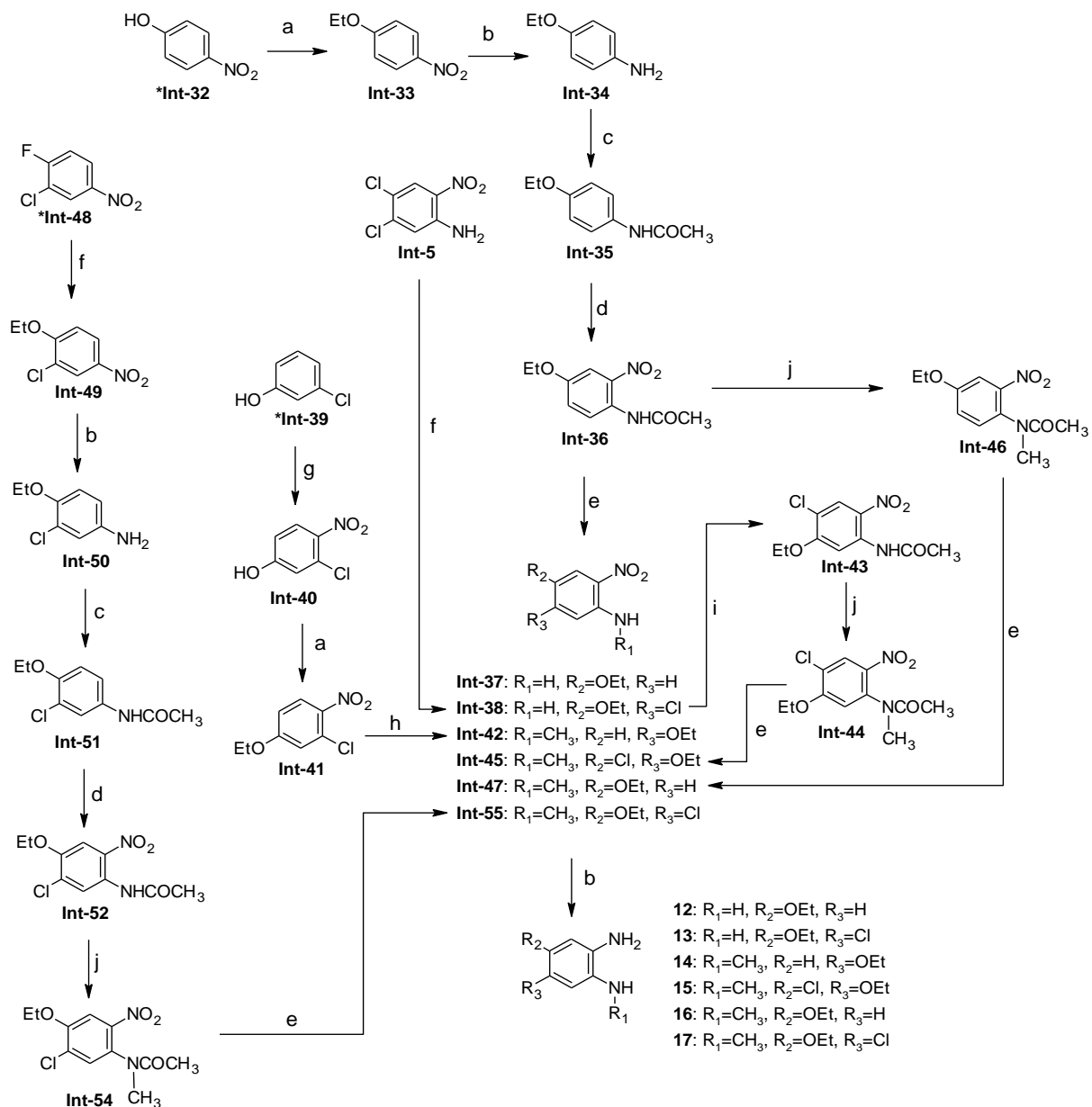


(a) MeOH, H₂SO₄, reflujo; (b) (CH₃)₂SO₄, K₂CO₃, acetona-H₂O, 40 °C; (c) KOH, MeOH; (d) CH₃NH₂·HCl, H₂O, K₂CO₃, 105 °C; (e) H₂, Pd/C (para compuestos no clorados) o Ni-Raney (para compuestos clorados); (f) (CH₃)₂SO₄, NaHCO₃, DMF, 65 °C, 2h; (g) Ac₂O; (h) H₂SO₄, HNO₃, 0 a 10 °C; (i) H₂SO₄, MeOH, H₂O (1.2 eq); (j) Ac₂O, H₂SO₄ (cat.); (k) (CH₃)₂SO₄, NaOH, H₂O, monoglina; (l) CH₃NH₂·HCl, DMF-H₂O, K₂CO₃, 130 °C.

Esquema 4. Síntesis de los 1,2-diaminobenzenos **6–11** a partir de compuestos comerciales (*).

6.4.2.4. 1,2-Diaminobencenos 12–17 (Serie 3)

La síntesis de los 1,2-diaminobencenos 12–17 se realizó mediante las rutas mostradas en el Esquema 5. Para la preparación de estos compuestos se partió de fenoles adecuadamente sustituidos o de compuestos halogenados que por sustitución nucleofílica dieron los derivados etoxi sustituidos.



(a) $CH_3CH_2I, K_2CO_3, DMF-H_2O, \Delta$; (b) $H_2, Pd/C$ (para compuestos no clorados) o $Ni-Raney$ (para compuestos clorados); (c) Ac_2O ; (d) $HNO_3, AcOH-H_2O, 10$ a $40\text{ }^\circ C$; (e) KOH o H_2SO_4, H_2O (1.2 eq), $MeOH, \Delta$; (f) $EtOH, Na^+, \Delta$; (g) $HNO_3-AcOH, -5$ a $0\text{ }^\circ C$; (h) $CH_3NH_2 \cdot HCl, DMF-H_2O, K_2CO_3, 110\text{ }^\circ C$; (i) Ac_2O, H_2SO_4 (cat.); (j) $(CH_3)_2SO_4, NaOH, H_2O$, monoglima.

Esquema 5. Síntesis de los 1,2-diaminobencenos 12–17 a partir de compuestos comerciales (*).

7. RESULTADOS Y DISCUSIÓN

7.1. Panorama de actividad tricomonocida y giardicida de derivados de bencimidazol

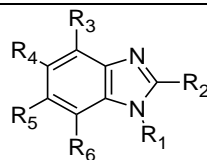
El estudio de las relaciones estructura-actividad juega un papel muy importante en la optimización de moléculas líder. Para estos estudios existen diferentes metodologías como los métodos basados en reglas, redes neuronales, modelado por farmacóforo, QSAR, entre otros [22, 24, 32–33]. Sin embargo, al aplicar muchas metodologías se asumen criterios que no son necesariamente válidos y esto puede llevar a resultados engañosos y modelos con poca o nula capacidad predictiva [72]. Por ejemplo, generalmente se asume que todos los compuestos en la base de datos tienen un modo de unión común y el mismo mecanismo de acción [27, 62]. Dado a que los estudios del panorama de actividad permiten caracterizar el SAR e identificar las características farmacofóricas, es crucial entender el panorama de actividad para obtener buenos modelos [52, 53].

Generalmente se asume, basándose en el principio de similitud, que moléculas similares tienen actividades similares; sin embargo, esto no es universal y es común encontrar bases de datos con relaciones estructura-actividad heterogéneas [52]. En este último caso se ha observado que es común encontrar moléculas donde pequeños cambios estructurales dirigen a cambios dramáticos en la actividad biológica. A este tipo de compuestos se les llama *activity cliffs* [52–53]. Cabe mencionar que los modelos QSAR convencionales no pueden explicar el comportamiento de los *activity cliffs*, y son generalmente vistos como *outliers*, por lo que es importante detectarlos antes de iniciar cualquier estudio [52]. Los compuestos cuya actividad biológica no puede ser predicha por modelos QSAR también han sido asociados con distintos modos de unión, mecanismos de acción diferentes y errores experimentales [53].

Al analizar el panorama de actividad también se pueden detectar moléculas estructuralmente distintas que tienen actividad similar (*Scaffold hopping*). Estas moléculas son relevantes porque ofrecen la posibilidad de mejorar propiedades fisicoquímicas y farmacocinéticas o acceder a nuevos compuestos con fines de propiedad intelectual [68].

En el presente trabajo se analizó el panorama de actividad para los derivados de bencimidazol presentados en la Tabla 2.

Tabla 2. Estructura, actividad tricomonocida y giardicida de derivados de bencimidazol usados en el estudio del panorama de actividad



	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	pCl ₅₀ <i>T. vaginalis</i> ^a	pCl ₅₀ <i>G. intestinalis</i> ^a	ΔpCl ₅₀
1	H	CF ₃	H	H	H	H	5.50	6.97	-1.47
2	CH ₃	CF ₃	H	CF ₃	H	H	5.39	5.94	-0.55
3	CH ₃	CF ₃	H	H	CF ₃	H	5.27	5.05	0.22
4	CH ₃	CF ₃	H	Propiltio	H	H	6.70	4.98	1.72
5	CH ₃	CF ₃	H	H	Propiltio	H	5.59	5.85	-0.26
6	CH ₃	CF ₃	H	Benzoilo	H	H	4.53	5.96	-1.43
7	CH ₃	CF ₃	H	H	Benzoilo	H	4.97	5.89	-0.92
8	H	CF ₃	H	Br	Br	H	6.66	6.92	-0.26
9	H	CF ₃	Br	Br	Br	Br	8.70	7.25	1.45
10	H	C ₂ F ₅	H	Cl	Cl	H	6.52	6.25	0.27
11	H	CF ₃	H	NO ₂	NO ₂	H	6.24	6.62	-0.38
12	H	C ₂ F ₅	Br	Br	Br	Br	5.00	7.64	-2.64
13	CH ₃	CONH ₂	H	H	Cl	H	6.96	7.12	-0.16
14	CH ₃	CONHCH ₃	H	H	Cl	H	6.98	7.15	-0.17
15	CH ₃	CON(CH ₃) ₂	H	H	Cl	H	6.63	7.40	-0.77
16	CH ₃	COOCH ₂ CH ₃	H	H	Cl	H	7.72	7.32	0.4
17	CH ₃	CONH ₂	H	Cl	H	H	6.73	6.63	0.1
18	CH ₃	CONHCH ₃	H	Cl	H	H	6.45	6.45	0
19	CH ₃	CON(CH ₃) ₂	H	Cl	H	H	6.68	6.61	0.07
20	CH ₃	COOCH ₂ CH ₃	H	Cl	H	H	7.57	7.40	0.17
21	CH ₃	CONH ₂	H	Cl	Cl	H	6.87	6.34	0.53
22	CH ₃	CONHCH ₃	H	Cl	Cl	H	6.65	6.82	-0.17
23	CH ₃	CON(CH ₃) ₂	H	Cl	Cl	H	7.12	7.13	-0.01
24	CH ₃	COOCH ₂ CH ₃	H	Cl	Cl	H	7.53	7.56	-0.03
25	CH ₃	CONH ₂	H	H	H	H	6.78	7.03	-0.25
26	CH ₃	CONHCH ₃	H	H	H	H	6.98	7.22	-0.24
27	CH ₃	CON(CH ₃) ₂	H	H	H	H	6.37	6.29	0.08
28	CH ₃	COOCH ₂ CH ₃	H	H	H	H	7.07	7.16	-0.09
29	CH ₃	COCH ₃	H	H	H	H	6.68	7.06	-0.38
30	CH ₃	COCH ₃	H	Cl	H	H	6.88	7.30	-0.42
31	CH ₃	COCH ₃	H	H	Cl	H	6.64	7.17	-0.53
32	CH ₃	COCH ₃	H	Cl	Cl	H	7.20	7.46	-0.26
33	H	CF ₃	H	Cl	H	H	6.35	5.89	0.46
34	H	CF ₃	H	H	Cl	H	6.35	5.89	0.46
35	H	CF ₃	H	F	H	H	5.50	6.31	-0.81
36	H	CF ₃	H	H	F	H	5.50	6.31	-0.81
37	H	CF ₃	H	CF ₃	H	H	6.63	6.17	0.46
38	H	CF ₃	H	H	CF ₃	H	6.63	6.17	0.46
39	H	CF ₃	H	CN	H	H	5.64	5.75	-0.11
40	H	CF ₃	H	H	CN	H	5.64	5.75	-0.11
41	H	CF ₃	H	Propiltio	H	H	6.46	5.82	0.64
42	H	CF ₃	H	H	Propiltio	H	6.46	5.82	0.64
43	H	CF ₃	H	Benzoilo	H	H	4.55	5.64	-1.09
44	H	CF ₃	H	H	Benzoilo	H	4.55	5.64	-1.09
45	H	CF ₃	H	Br	H	H	5.80	6.19	-0.39

46	H	CF ₃	H	H	Br	H	5.80	6.19	-0.39
47	H	CF ₃	Br	H	Br	H	6.72	6.60	0.12
48	H	CF ₃	H	Br	H	Br	6.72	6.60	0.12
49	H	CF ₃	Br	Br	Br	H	6.57	6.85	-0.28
50	H	CF ₃	H	Br	Br	Br	6.57	6.85	-0.28
51	H	SCH ₂ CH ₂ OH	Cl	H	Cl	H	6.16	7.42	-1.26
52	H	SCH ₂ CH ₂ OH	H	Cl	H	Cl	6.16	7.42	-1.26
53	H	SCH ₂ CH ₂ OH	Br	H	Br	H	6.83	6.98	-0.15
54	H	SCH ₂ CH ₂ OH	H	Br	H	Br	6.83	6.98	-0.15
55	H	SCH ₂ CH ₂ N(CH ₃) ₂	Cl	H	Cl	H	7.14	7.68	-0.54
56	H	SCH ₂ CH ₂ N(CH ₃) ₂	H	Cl	H	Cl	7.14	7.68	-0.54
57	H	SCH ₂ CH ₂ N(CH ₃) ₂	Br	H	Br	H	6.51	7.54	-1.03
58	H	SCH ₂ CH ₂ N(CH ₃) ₂	H	Br	H	Br	6.51	7.54	-1.03
59	H	SCH ₂ CH ₂ N(C ₂ H ₅) ₂	Cl	H	Cl	H	7.39	7.72	-0.33
60	H	SCH ₂ CH ₂ N(C ₂ H ₅) ₂	H	Cl	H	Cl	7.39	7.72	-0.33
61	H	SCH ₂ CH ₂ N(C ₂ H ₅) ₂	Br	H	Br	H	7.53	7.46	0.07
62	H	SCH ₂ CH ₂ N(C ₂ H ₅) ₂	H	Br	H	Br	7.53	7.46	0.07
63	H	SCH ₂ CH ₂ CH ₂ N(CH ₃) ₂	Cl	H	Cl	H	7.07	7.14	-0.07
64	H	SCH ₂ CH ₂ CH ₂ N(CH ₃) ₂	H	Cl	H	Cl	7.07	7.14	-0.07
65	H	SCH ₂ CH ₂ CH ₂ N(CH ₃) ₂	Br	H	Br	H	7.99	7.69	0.3
66	H	SCH ₂ CH ₂ CH ₂ N(CH ₃) ₂	H	Br	H	Br	7.99	7.69	0.3
67	H	SCH ₂ CH ₂ -(<i>N</i> -piperidil)	Cl	H	Cl	H	7.14	7.26	-0.12
68	H	SCH ₂ CH ₂ -(<i>N</i> -piperidil)	H	Cl	H	Cl	7.14	7.26	-0.12
69	H	SCH ₂ CH ₂ -(<i>N</i> -piperidil)	Br	H	Br	H	7.69	6.90	0.79
70	H	SCH ₂ CH ₂ -(<i>N</i> -piperidil)	H	Br	H	Br	7.69	6.90	0.79
71	H	SCH ₂ CH ₂ -(<i>N</i> -morfolinil)	Cl	H	Cl	H	7.29	7.41	-0.12
72	H	SCH ₂ CH ₂ -(<i>N</i> -morfolinil)	H	Cl	H	Cl	7.29	7.41	-0.12
73	H	SCH ₂ CH ₂ -(<i>N</i> -morfolinil)	Br	H	Br	H	7.23	7.23	0
74	H	SCH ₂ CH ₂ -(<i>N</i> -morfolinil)	H	Br	H	Br	7.23	7.23	0
75	H	SCH ₂ CH ₂ -(<i>p</i> -nitrofenil)	Cl	H	Cl	H	7.96	7.09	0.87
76	H	SCH ₂ CH ₂ -(<i>p</i> -nitrofenil)	H	Cl	H	Cl	7.96	7.09	0.87
77	H	SCH ₂ CH ₂ -(<i>p</i> -nitrofenil)	Br	H	Br	H	8.48	7.43	1.05
78	H	SCH ₂ CH ₂ -(<i>p</i> -nitrofenil)	H	Br	H	Br	8.48	7.43	1.05

^aLos datos de actividad se expresan en $-\log\text{Cl}_{50} = \text{pCl}_{50}$.

El trabajo comprende dos secciones considerando distintos objetivos y tomando en cuenta la naturaleza de las moléculas en la base de datos. En la primera sección se buscó la caracterización del panorama de actividad para 32 moléculas (1–32), las cuales no poseen formas tautoméricas biológicamente importantes. En una segunda aproximación se consideraron 55 moléculas de la base de datos con el fin de vincular las relaciones estructura-actividad para dos protozoarios. Cabe mencionar que debido a que 23 moléculas tienen dos tautómeros relevantes, en este estudio se consideraron 78 estructuras que incluyen formas tautoméricas (1–78)

7.1.1. Estudio basado en 32 compuestos

7.1.1.1. Representaciones de similitud estructural

Dado a que diferentes medidas de similitud molecular capturan diferentes aspectos de la estructura molecular [73–74], el panorama de actividad es muy dependiente de la representación usada. En un esfuerzo por mejorar en este aspecto y capturar la mayor información posible en una medida de similitud molecular se usaron distintas representaciones no correlacionadas y se generaron modelos consenso [55, 75]. La Tabla 3 muestra las diferentes representaciones moleculares calculadas en este estudio así como una breve descripción de ellas.

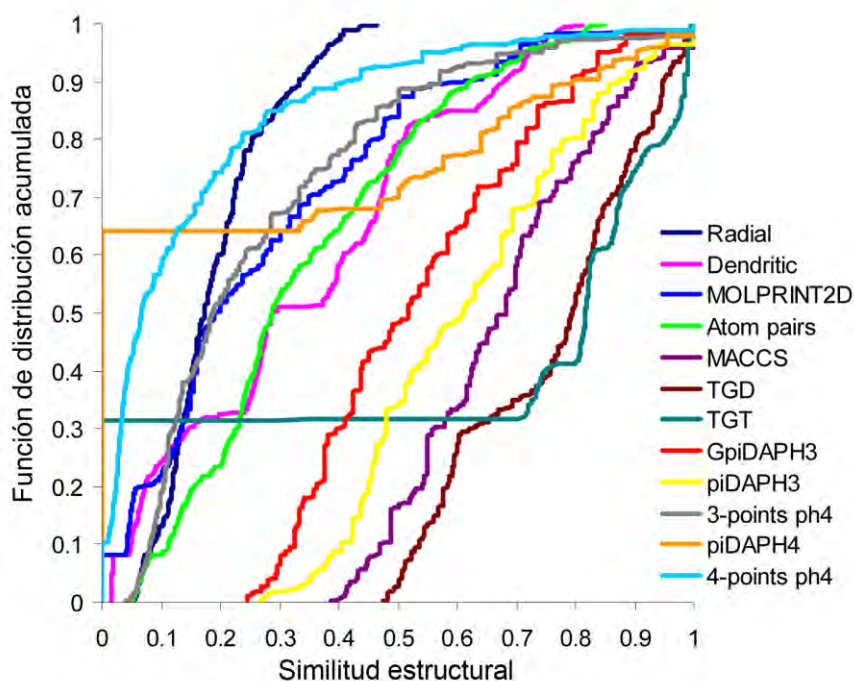
Tabla 3. Representaciones moleculares calculadas

Representación	Descripción	
2D	MACCS keys (166 bits)	Basada en comparaciones de subestructuras
	GpiDAPH3	Basada en triángulos de farmacóforo en dos dimensiones
	TGD	Basada distancias gráficas entre dos átomos
	TGT	Basada distancias gráficas entre tres átomos
	Radial	Basada en fragmentos extendidos radialmente para cada átomo
	Dendritic	Basada en fragmentos lineales y ramificados obtenidos mediante descomposición de la estructura
	Atom pairs	Basada en la diferenciación de tipos de átomos y la distancia entre ellos
	MOLPRINT 2D	Variante de la representación radial que integra tipos de átomos localizados a diferentes distancias
	Propiedades	Basada en propiedades calculadas a partir de la estructura molecular: peso molecular, número de enlaces rotables, número de aceptores de hidrógeno, número de donadores de hidrógeno, Área de superficie polar y coeficiente de partición octanol/agua.
3D	piDAPH3	Basada en 3 puntos de farmacóforo
	piDAPH4	Basada en 4 puntos de farmacóforo
	3-point pharmacophore	Basada en 3 puntos de farmacóforo
	4-point pharmacophore	Basada en 4 puntos de farmacóforo

7.1.1.2. Distribución de las representaciones de similitud estructural

Para los 32 compuestos considerados en esta sección (1–32, Tabla 2), se calcularon 496 comparaciones pareadas de similitud estructural usando 12 metodologías de cálculo (sección 6.1.3). Posteriormente se calcularon las curvas de

distribución acumulada y los valores máximo (Max), tercer cuartil (Q3), mediana, primer cuartil (Q1), mínimo (Min), promedio y desviación estándar (σ).



	Max	Q3	Mediana	Q1	Min	Promedio	σ
Radial	0.46	0.24	0.17	0.13	0.05	0.19	0.09
Dendritic	0.81	0.48	0.29	0.11	0.02	0.33	0.23
MOLPRINT 2D	1.00	0.42	0.20	0.12	0.00	0.27	0.22
Atom pairs	0.85	0.49	0.29	0.21	0.05	0.34	0.20
MACCS	1.00	0.80	0.68	0.55	0.39	0.68	0.16
TGD	1.00	0.88	0.80	0.59	0.47	0.76	0.16
TGT	1.00	0.90	0.82	0.00	0.00	0.60	0.41
GpiDAPH3	1.00	0.69	0.52	0.38	0.25	0.54	0.18
piDAPH3	1.00	0.76	0.61	0.46	0.27	0.62	0.18
3-points ph4	1.00	0.37	0.19	0.11	0.04	0.27	0.21
piDAPH4	1.00	0.58	0.00	0.00	0.00	0.24	0.34
4-points ph4	1.00	0.20	0.07	0.03	0.00	0.15	0.20

Figura 15. Funciones de distribución acumulada para distintas representaciones de similitud estructural en 2D y 3D. La tabla resume la información de las distribuciones. Q3 y Q1 representan el tercer y primer cuartil, respectivamente; σ representa la desviación estándar.

Las diferentes representaciones abarcan un amplio rango de variación en su distribución. Las representaciones en 2D con los valores más altos de similitud y desviación estándar comparable son TGD, MACCS y GpiDAPH3 con valores para la mediana de 0.8, 0.68 y 0.52, respectivamente. Las representaciones 2D con menores

valores de similitud fueron dendritic, atom pairs, MOLPRINT 2D y radial. En lo que respecta a las representaciones en 3D, la representación basada en tres puntos de farmacóforo piDAPH3 tiene una distribución caracterizada por una sigmoide. En contraste, las representaciones piDAPH4, 3- y 4-point pharmacophores tienen una distribución caracterizada por una forma no sigmoidea de las funciones de distribución. Aunque todas las moléculas tienen en común el núcleo de bencimidazol y en muchos casos las variaciones en los sustituyentes son pequeñas, varias representaciones son capaces de diferenciar las moléculas presentes en esta base de datos.

7.1.1.3. Correlación entre similitudes moleculares

Con el fin de identificar las representaciones moleculares menos correlacionadas se calculó la matriz de correlación usando 496 similitudes pareadas para cada una de las 12 representaciones de similitud estructural, la similitud en propiedades y la similitud en actividad (Tabla 4).

Tabla 4. Correlación entre diferentes representaciones de similitud molecular

	Radial	Dendritic	MOLPRINT 2D	Atom pairs	MACCS	TGD	TGT	GpiDAPH3	piDAPH3	3-points ph4	piDAPH4	4-points ph4	Propiedades	SA ^a (T. vaginalis)	SA ^a (G. intestinalis)
Radial	1.00														
Dendritic	0.94	1.00													
MOLPRINT 2D	0.88	0.86	1.00												
Atom pairs	0.92	0.92	0.93	1.00											
MACCS	0.78	0.81	0.84	0.89	1.00										
TGD	0.71	0.82	0.69	0.76	0.77	1.00									
TGT	0.69	0.82	0.63	0.71	0.73	0.93	1.00								
GpiDAPH3	0.79	0.76	0.77	0.83	0.76	0.69	0.56	1.00							
piDAPH3	0.76	0.73	0.74	0.79	0.73	0.69	0.57	0.94	1.00						
3-points ph4	0.78	0.75	0.81	0.83	0.77	0.68	0.55	0.82	0.81	1.00					
piDAPH4	0.40	0.36	0.43	0.48	0.53	0.32	0.19	0.61	0.62	0.49	1.00				
4-points ph4	0.73	0.69	0.78	0.79	0.72	0.63	0.49	0.78	0.77	0.97	0.50	1.00			
Propiedades	0.69	0.69	0.74	0.78	0.78	0.65	0.59	0.73	0.74	0.75	0.44	0.69	1.00		
SA ^a (T. vaginalis)	0.43	0.40	0.44	0.45	0.28	0.26	0.19	0.48	0.47	0.42	0.25	0.39	0.37	1.00	
SA ^a (G. intestinalis)	0.28	0.20	0.28	0.25	0.15	-0.01	-0.05	0.29	0.30	0.27	0.17	0.26	0.20	0.31	1.00

^aSA: Similitud en actividad

Se observan correlaciones altas entre muchas de las huellas digitales en 2D, por ejemplo, MOLPRINT 2D y atom pairs; dendritic y atom pairs; TGD y TGT (correlación ≥ 0.92). Se observan otras correlaciones altas para los métodos en 2D atom pairs y MACCS (0.89); MOLPRINT 2D y MACCS (0.84); atom pairs y GpiDAPH3 (0.83). También se encontró correlación alta para huellas digitales en 3D como 3- y 4-points pharmacophore (0.97), y piDAPH3 y 3-points pharmacophore (0.82). Cuando se comparan las huellas digitales 2D y 3D la correlación más alta es para GpiDAPH3 y piDAPH3 (0.94). La matriz muestra baja a moderada correlación para las huellas digitales y la similitud en propiedades (0.44–0.78). Por otra parte, las huellas digitales muestran baja correlación con la similitud en actividad para *T. vaginalis* (correlación ≤ 0.48) y *G. intestinalis* (correlación ≤ 0.30). La correlación entre la similitud en actividad para ambos parásitos también es baja (0.31), indicando la presencia de compuestos con diferente efecto contra ambos parásitos.

Aunque las correlaciones entre las huellas digitales (2D y 3D) son en general altas, se escogieron las representaciones menos correlacionadas para el estudio del panorama de actividad. En el caso de las representaciones 2D se seleccionaron las huellas digitales radial, MACCS y TGD; para las representaciones en 3D se seleccionó piDAPH3. Se incluyó también la similitud basada en propiedades. La correlación máxima entre estas representaciones fue 0.78 (radial y MACCS) y la mínima 0.65 (propiedades y MACCS).

7.1.1.4. Mapas SAS

En las Figuras 16 y 17 se muestran los mapas SAS para *T. vaginalis* y *G. intestinalis*, respectivamente. Los mapas muestran la relación entre la similitud en actividad y la similitud molecular obtenida por las huellas digitales seleccionadas y las propiedades calculadas. Cada gráfico contiene 496 puntos representando una comparación pareada cada uno. Adicionalmente, cada punto se codificó en una escala continua de color para distinguir aquellos pares con al menos un compuesto activo (rojos) de los pares donde ambos compuestos son inactivos (verdes). Como se esperaba, según la información de las funciones de distribución acumulada (Figura 15), los valores de similitud obtenidos para la huella digital radial son bajos y los

puntos se encuentran acumulados en la parte izquierda de la gráfica ($X < 0.46$) (Figura 16A); mientras que en los gráficos obtenidos con MACCS, TGD y piDAPH3 muestran puntos con valores más altos y más dispersos. Se observó que varios puntos en las Figuras 16 y 17 toman valores de 1 (21 pares en MACCS, 15 en TGD, 18 en piDAPH3 y 8 en similitud en propiedades); Muchos de esos puntos corresponden a isómeros constitucionales o compuestos sustituidos con diferentes halógenos. No obstante, la huella digital radial distingue los 496 pares. Este comportamiento es esperado dado a que la huella digital radial fue diseñada para estudios estructura–actividad, mientras que MACCS, TGD y piDAPH3 fueron desarrollados para búsquedas virtuales por similitud [31, 76].

Como se describió en las secciones 2.3.3 (Figura 14) y 6.1.6, los mapas se dividieron en cuatro regiones (I–IV). Los puntos localizados en las regiones I y II tienen relaciones estructura–actividad continuas, mientras que los pares en la región IV representan *activity cliffs*. Como se mencionó previamente, la división entre las regiones I/II y III/IV depende de la representación usada; sin embargo, se pueden detectar los pares situados en la misma región relativa para diferentes mapas (pares consenso). En las Figuras 16 y 17 se marcan algunos ejemplos de pares consenso para todas las representaciones. El par **8_28** en la región I y el par **16_24** en la región II representan pares consenso para *T. vaginalis* y *G. intestinalis*, mientras que **9_12** es un par consenso en la región IV para *T. vaginalis*.

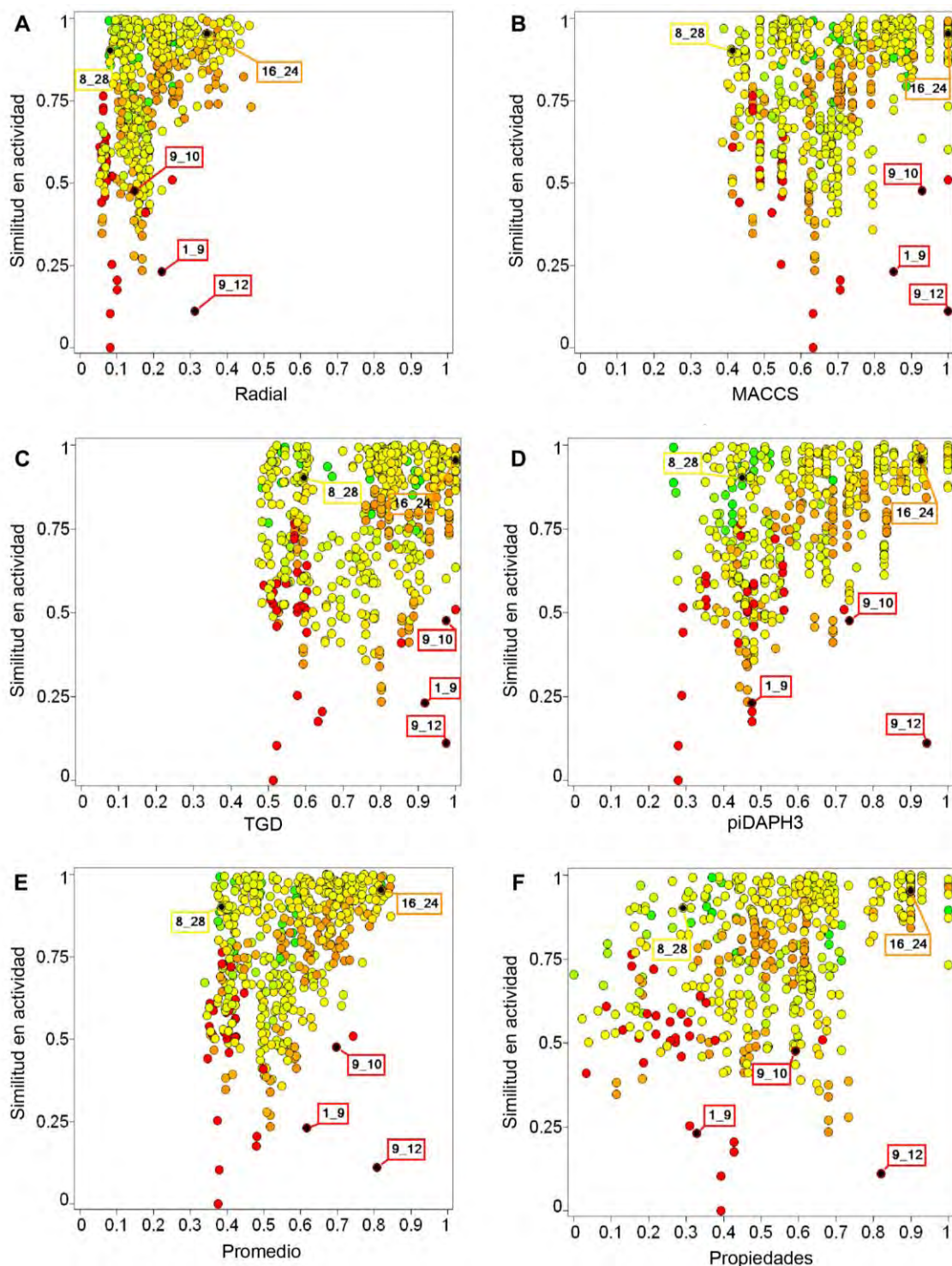


Figura 16. Mapas SAS para *T. vaginalis* con diferentes representaciones moleculares. Cada punto significa una comparación pareada de los 32 bencimidazoles estudiados (496 puntos). Los puntos están codificados en escala continua de color siendo rojos los pares más activos (al menos un compuesto en el par tiene $pCl_{50} \geq 7$) y verdes los pares menos activos (ambos compuestos tienen $pCl_{50} < 7$). Cada panel muestra una representación molecular diferente: (A) Radial; (B) MACCS; (C) TGD; (D) π iDAPH3; (E) promedio de huellas digitales; (F) Propiedades. Los pares seleccionados están marcados en negro y etiquetados con los números correspondientes a los compuestos presentes.

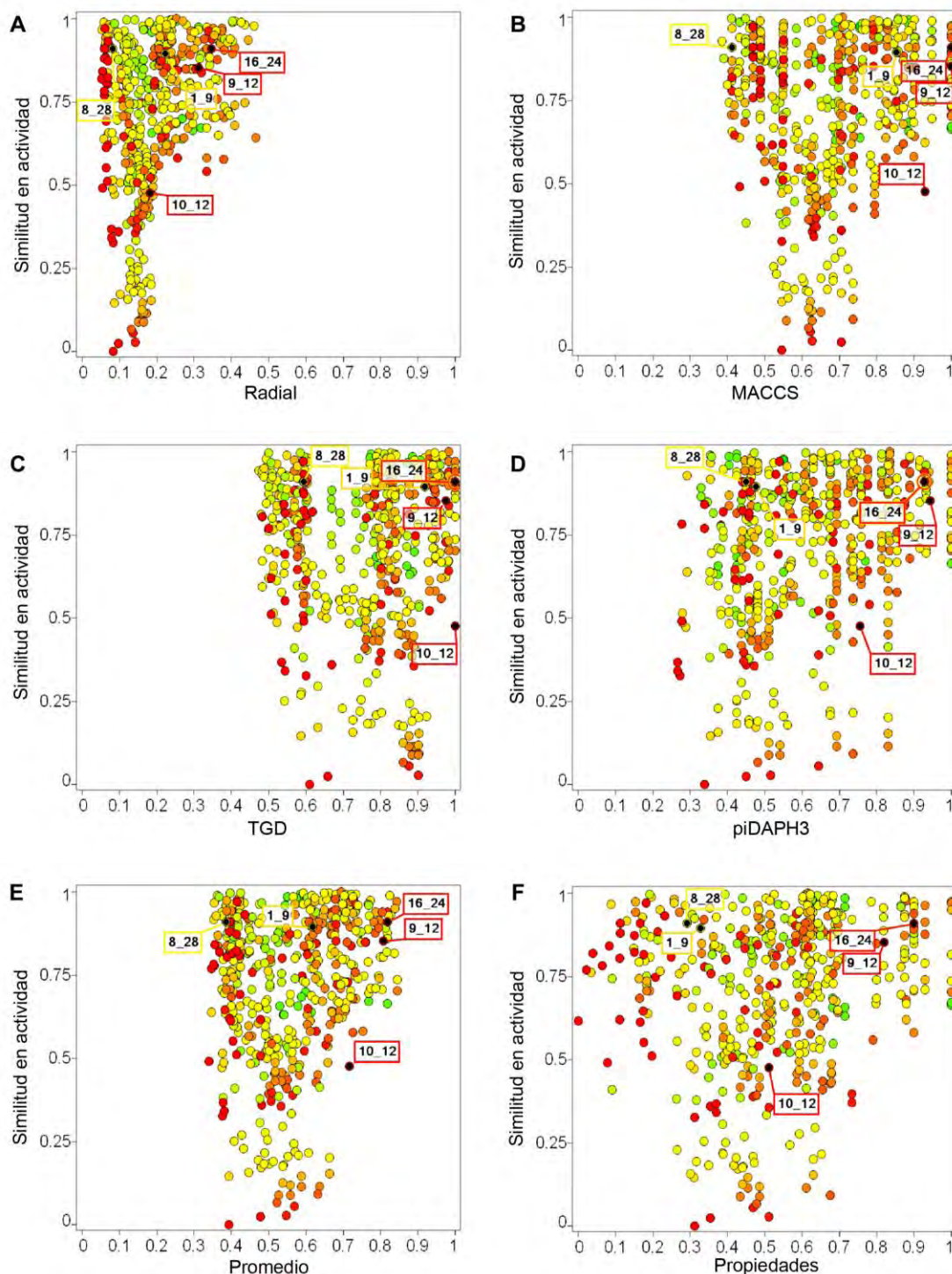


Figura 17. Mapas SAS para *G. intestinalis* con diferentes representaciones moleculares. Cada punto significa una comparación pareada de los 32 bencimidazoles estudiados (496 puntos). Los puntos están codificados en escala continua de color donde los pares más activos se observan en rojo (al menos un compuesto en el par tiene $pCl_{50} \geq 7.3$) y los menos activos en verde (ambos compuestos tienen $pCl_{50} < 7.3$). Cada panel muestra una representación molecular diferente: (A) Radial; (B) MACCS; (C) TGD; (D) piDAPH3; (E) promedio de huellas digitales; (F) Propiedades. Los pares seleccionados están marcados en negro y etiquetados con los números correspondientes a los compuestos presentes.

7.1.1.5. Caracterización cuantitativa de los mapas SAS

En esta sección se describe la cuantificación de los pares en cada región (I–IV) de los mapas SAS generados con diferentes representaciones de similitud molecular (Tabla 5). Las cuatro regiones se dividieron mediante la metodología descrita en la sección 6.1.6. Para delimitar la región de similitud estructural se tomaron en cuenta los siete compuestos más activos para ambos parásitos (mediana de la similitud de los siete compuestos más activos). En la Tabla 5 también se indica el número de pares donde al menos uno de los compuestos tiene actividad alta (valores definidos como ≥ 7.0 para *T. vaginalis* y ≥ 7.3 para *G. intestinalis*).

Tabla 5. Distribución de datos en cada región de los mapas SAS

Representación (SA límite) ^a	Mediana de los activos ^b	I		II		III		IV	
		Pares Totales ^c	Pares activos ^d	Pares Totales	Pares activos	Pares Totales	Pares activos	Pares Totales	Pares activos
<i>T. vaginalis</i> (0.5)									
Radial	0.22	286	94	160	66	47	34	3	2
MACCS	0.71	271	102	175	58	42	31	8	5
TGD	0.95	401	138	45	22	48	34	2	2
piDAPH3	0.78	346	129	100	31	49	35	1	1
Propiedades	0.62	285	120	161	40	38	25	12	11
<i>T. vaginalis</i> (0.75)									
Radial	0.22	145	50	149	59	188	78	14	9
MACCS	0.71	142	59	152	50	171	74	31	13
TGD	0.95	254	89	40	20	195	83	7	4
piDAPH3	0.78	204	80	90	29	191	84	11	3
Propiedades	0.62	160	74	134	35	163	71	39	16
<i>G. intestinalis</i> (0.5)									
Radial	0.24	301	83	114	50	81	38	0	0
MACCS	0.79	286	91	129	42	76	35	5	3
TGD	0.93	327	90	88	43	80	37	1	1
piDAPH3	0.84	356	113	59	20	81	38	0	0
Propiedades	0.61	241	87	174	46	63	26	18	12
<i>G. intestinalis</i> (0.75)									
Radial	0.24	184	54	76	35	198	67	38	15
MACCS	0.79	171	58	89	31	191	68	45	14
TGD	0.93	199	57	61	32	208	70	28	12
piDAPH3	0.84	216	72	44	17	221	79	15	3
Propiedades	0.61	138	54	122	35	166	59	70	23

^aLas regiones I–IV están definidas por la mediana de la similitud para los compuestos activos y 0.5 ó 0.75 como división para similitud en actividad (ver también la Figura 14 y el texto para detalles).

^bMediana de la similitud molecular para los compuestos **9, 16, 20, 23, 24, 28, 32** para *T. vaginalis*, y **12, 15, 16, 20, 24, 30, 32** para *G. intestinalis*.

^cNúmero total de puntos (par de compuestos) en la región.

^dNúmero total de puntos con al menos un compuesto activo en el par.

Otra manera de visualizar los datos en cada región es usando gráficos circulares. En la Figura 18 se muestra un ejemplo para *T. vaginalis* usando un valor límite de 0.5 para la similitud en actividad.

Considerando la división de 0.5 para similitud en actividad, para ambos protozoarios, el mayor número de puntos se ubica sobre la región I y la frecuencia disminuye para la mayoría de las representaciones en el orden I > II > III > IV (Tabla 5).

Es de esperarse que para la división de 0.75 de similitud de actividad el número total de pares, y los pares activos, disminuya en las regiones I y II y aumente en las regiones III y IV cuando se compara con una división de 0.5. En este caso, la región con mayor número de pares es la III en todas las representaciones para *G. intestinalis* y en la mayoría de los casos para *T. vaginalis*. Los resultados son similares para los pares activos. En general, se observa la tendencia en el número de pares en cada región en el orden III > I > II > IV (Tabla 5).

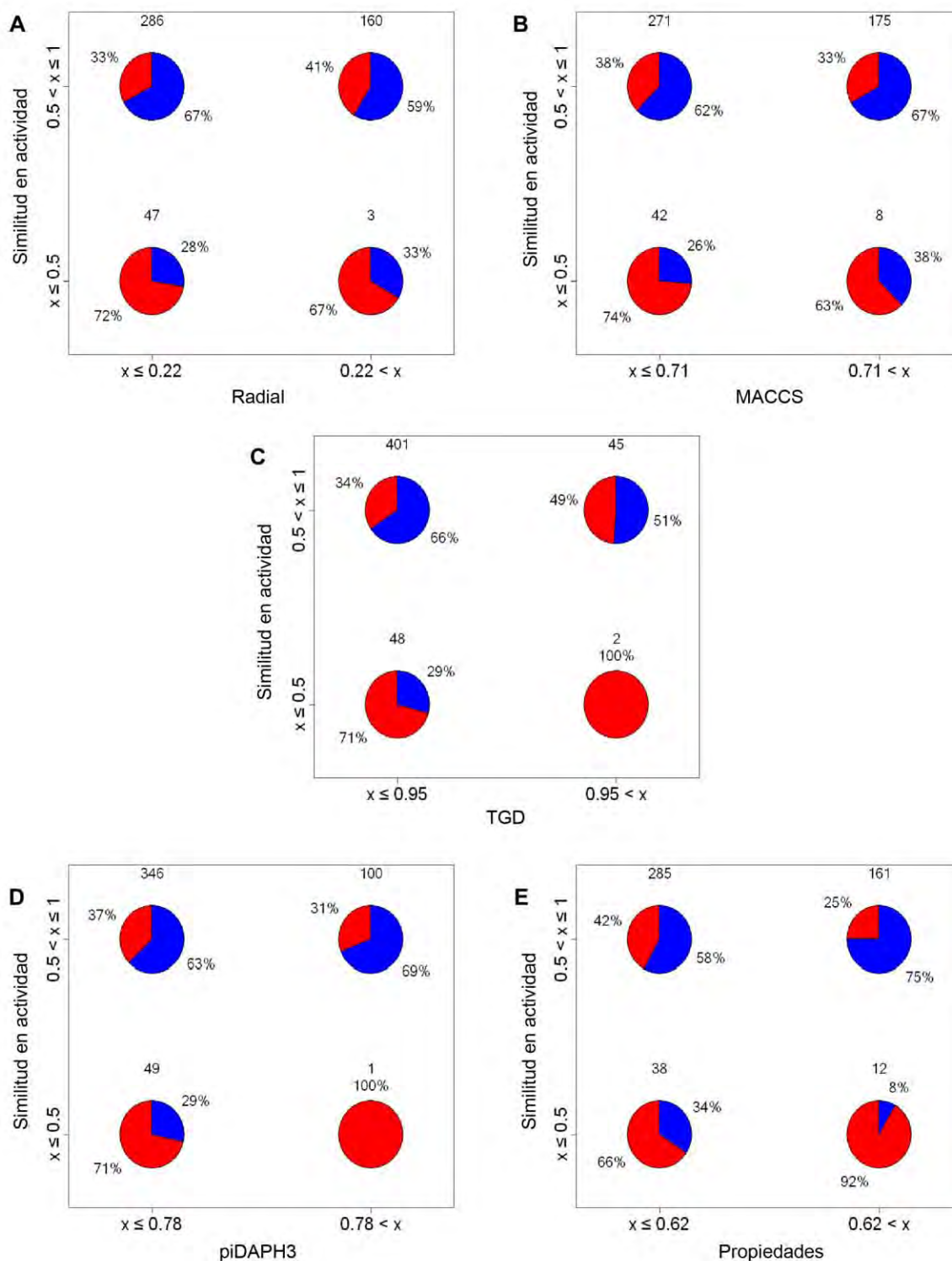


Figura 18. Ejemplos de gráficos circulares para *T. vaginalis* obtenidos con una división en similitud de actividad de 0.5 y diferentes representaciones moleculares: (A) Radial; (B) MACCS; (C) piDAPH3; (E) Propiedades. Los gráficos indican el número de puntos en cada una de las cuatro regiones definidas en la Figura 14. En color rojo se indica la proporción de pares con al menos un compuesto activo ($pCl_{50} \geq 7.0$; $Cl_{50} \leq 100$ nM).

7.1.1.6. *Activity cliffs* profundos y superficiales

Los *activity cliffs* pueden ser clasificados tomando en cuenta diferentes valores para el límite de similitud en actividad. En este trabajo se define como *activity cliffs profundos* a pares de compuestos similares que tienen alta diferencia en actividad (similitud en actividad ≤ 0.5); del mismo modo se define como *activity cliffs superficiales* los pares con diferencia en actividad menor ($0.5 < \text{similitud en actividad} \leq 0.75$) [77].

De acuerdo a las definiciones anteriores encontramos entre uno y ocho *activity cliffs profundos* para *T. vaginalis* dependiendo de la huella digital empleada. El número de *activity cliffs profundos* para *G. intestinalis* es menor, entre cero y cinco, dependiendo de la representación empleada. Considerando las propiedades como representación, se encontraron 12 y 18 *activity cliffs profundos* para *T. vaginalis* y *G. intestinalis*, respectivamente. Es importante aclarar que el número de *activity cliffs profundos* y *superficiales* es relativo a estos 32 compuestos estudiados, ya que depende del rango de actividad en la base de datos.

El número de *activity cliffs superficiales* puede ser calculado usando los valores de la Tabla 5 mediante la diferencia de los pares totales en la región IV a los límites de 0.5 y 0.75 de similitud en actividad. Por ejemplo, para *T. vaginalis*, hay $14 - 3 = 11$ *activity cliffs superficiales* considerando la huella digital radial y $31 - 8 = 23$ considerando la huella digital MACCS. En la Tabla 6 se resume el número de *activity cliffs superficiales* para cada representación en ambos protozoarios. Es importante notar que el número de *activity cliffs superficiales* para *G. intestinalis* es mayor que para *T. vaginalis*.

Tabla 6. Número total de *activity cliffs superficiales* en diferentes representaciones moleculares

Representación Molecular	Pares en la región IV <i>T. vaginalis</i>			Pares en la región IV <i>G. intestinalis</i>		
	DSA ^a (0.75)	DSA (0.5)	<i>activity cliffs superficiales</i>	DSA (0.75)	DSA (0.5)	<i>activity cliffs superficiales</i>
Radial	14	3	11	38	0	38
MACCS	31	8	23	45	5	40
TGD	7	2	5	28	1	27
piDAPH3	11	1	10	15	0	15
Propiedades	39	12	27	70	18	52

^aDSA: División de similitud en actividad.

7.1.1.7. Grado de consenso

Debido a que diferentes representaciones de similitud estructural capturan diferentes características moleculares, el número de pares en la base de datos ubicados en cada región depende del criterio usado para definir los compuestos como similares. Aunque las representaciones moleculares seleccionadas muestran correlaciones que van desde moderadas a bajas, y una distribución diferente en los valores de similitud, es posible encontrar un número considerable de pares consenso en diferentes regiones de los mapas SAS. Se calculó el número de pares de compuestos que dos métodos ponen en la misma región (pares consenso) para ambos protozoarios a dos diferentes divisiones de similitud en actividad (Tabla 7). Se encontraron muchos pares consenso para ambos parásitos en todas las regiones. La única excepción fue para la región IV para *G. intestinalis* cuando se usó una división de similitud en actividad (DSA) de 0.5; sólo se encontró un par consenso entre MACCS y TGD, y dos pares consenso entre MACCS y similitud en propiedades (Tabla 7). El grado de consenso (Doc) entre dos métodos se presenta en la Tabla 8. El Doc mide el número de pares consenso entre dos métodos escalado con el número de pares totales que ambos métodos ponen en la misma región (ver la sección 6.1.6). Para ambos protozoarios, a ambos valores de DSA, Doc tiene valores altos en la región III (0.70–1.00) seguido por la región I (0.49–0.98). En contraste, Doc tiene valores bajos en la región IV, especialmente para *G. intestinalis* para el valor de DSA de 0.5 (0–0.2).

En la Tabla 9 se muestran varios ejemplos de pares consenso para varias representaciones en las tres regiones más relevantes (I, II y IV) para ambos protozoarios.

Tabla 7. Número de pares consenso en las cuatro regiones de los mapas SAS para las representaciones moleculares: Radial, MACCS, TGD, piDAPH3 y similitud en propiedades (SP). Los números en la diagonal indican la cantidad de pares que se encuentran en una representación en particular

T. vaginalis (división similitud en actividad 0.5)

Región I						Región II					
	Radial	MACCS	TGD	piDAPH3	SP		Radial	MACCS	TGD	piDAPH3	SP
Radial	286					Radial	160				
MACCS	250	271				MACCS	139	175			
TGD	275	258	401			TGD	34	32	45		
piDAPH3	275	261	333	346		piDAPH3	89	90	32	100	
SP	239	246	274	269	285	SP	114	136	34	84	161

Región III						Región IV					
	Radial	MACCS	TGD	piDAPH3	SP		Radial	MACCS	TGD	piDAPH3	SP
Radial	47					Radial	3				
MACCS	41	42				MACCS	2	8			
TGD	46	42	48			TGD	1	2	2		
piDAPH3	47	42	48	49		piDAPH3	1	1	1	1	
SP	37	33	37	38	38	SP	2	3	1	1	12

G. intestinalis (división similitud en actividad 0.5)

Región I						Región II					
	Radial	MACCS	TGD	piDAPH3	SP		Radial	MACCS	TGD	piDAPH3	SP
Radial	301					Radial	114				
MACCS	255	286				MACCS	83	129			
TGD	270	264	327			TGD	57	66	88		
piDAPH3	286	285	317	356		piDAPH3	44	58	49	59	
SP	217	229	230	241	241	SP	90	117	77	59	174

Región III						Región IV					
	Radial	MACCS	TGD	piDAPH3	SP		Radial	MACCS	TGD	piDAPH3	SP
Radial	81					Radial	0				
MACCS	76	76				MACCS	0	5			
TGD	80	76	80			TGD	0	1	1		
piDAPH3	81	76	80	81		piDAPH3	0	0	0	0	
SP	63	60	62	63	63	SP	0	2	0	0	18

T. vaginalis (división similitud en actividad 0.75)

Región I						Región II					
	Radial	MACCS	TGD	piDAPH3	SP		Radial	MACCS	TGD	piDAPH3	SP
Radial	145					Radial	149				
MACCS	122	142				MACCS	129	152			
TGD	136	130	254			TGD	31	28	40		
piDAPH3	142	137	195	204		piDAPH3	87	85	31	90	
SP	120	133	152	150	160	SP	109	125	32	80	134

Región III						Región IV					
	Radial	MACCS	TGD	piDAPH3	SP		Radial	MACCS	TGD	piDAPH3	SP
Radial	188					Radial	14				
MACCS	169	171				MACCS	12	31			
TGD	185	170	195			TGD	4	6	7		
piDAPH3	180	166	186	191		piDAPH3	3	6	2	11	
SP	156	146	159	157	163	SP	7	14	3	5	39

G. intestinalis (división similitud en actividad 0.75)

Región I						Región II					
	Radial	MACCS	TGD	piDAPH3	SP		Radial	MACCS	TGD	piDAPH3	SP
Radial	184					Radial	76				
MACCS	153	171				MACCS	58	89			
TGD	165	159	199			TGD	42	49	61		
piDAPH3	173	170	192	216		piDAPH3	33	43	37	44	
SP	123	132	133	138	138	SP	61	83	56	44	122

Región III						Región IV					
	Radial	MACCS	TGD	piDAPH3	SP		Radial	MACCS	TGD	piDAPH3	SP
Radial	198					Radial	38				
MACCS	178	191				MACCS	25	45			
TGD	185	181	208			TGD	15	18	28		
piDAPH3	194	191	205	221		piDAPH3	11	15	12	15	
SP	157	157	159	166	166	SP	29	36	21	15	70

Tabla 8. Grado de consenso (Doc) en las cuatro regiones de los mapas SAS para las representaciones moleculares: Radial, MACCS, TGD, piDAPH3 y similitud en propiedades (SP). Cada valor en las matrices representa el grado en que dos métodos colocan pares en la misma región en una escala de 0–1

T. vaginalis (división similitud en actividad 0.5)

Región I

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.81	1.00			
TGD	0.67	0.62	1.00		
piDAPH3	0.77	0.73	0.80	1.00	
SP	0.72	0.79	0.67	0.74	1.00

Región II

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.71	1.00			
TGD	0.20	0.17	1.00		
piDAPH3	0.52	0.49	0.28	1.00	
SP	0.55	0.68	0.20	0.47	1.00

Región III

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.85	1.00			
TGD	0.94	0.88	1.00		
piDAPH3	0.96	0.86	0.98	1.00	
SP	0.77	0.70	0.76	0.78	1.00

Región IV

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.22	1.00			
TGD	0.25	0.25	1.00		
piDAPH3	0.33	0.13	0.50	1.00	
SP	0.15	0.18	0.08	0.08	1.00

G. intestinalis (división similitud en actividad 0.5)

Región I

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.77	1.00			
TGD	0.75	0.76	1.00		
piDAPH3	0.77	0.80	0.87	1.00	
SP	0.67	0.77	0.68	0.68	1.00

Región II

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.52	1.00			
TGD	0.39	0.44	1.00		
piDAPH3	0.34	0.45	0.50	1.00	
SP	0.45	0.63	0.42	0.34	1.00

Región III

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.94	1.00			
TGD	0.99	0.95	1.00		
piDAPH3	1.00	0.94	0.99	1.00	
SP	0.78	0.76	0.77	0.78	1.00

Región IV

	Radial	MACCS	TGD	piDAPH3	SP
Radial	NA				
MACCS	0.00	1.00			
TGD	0.00	0.20	1.00		
piDAPH3	NA	0.00	0.00	NA	
SP	0.00	0.10	0.00	0.00	1.00

T. vaginalis (división similitud en actividad 0.75)

Región I

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.74	1.00			
TGD	0.52	0.49	1.00		
piDAPH3	0.69	0.66	0.74	1.00	
SP	0.65	0.79	0.58	0.70	1.00

Región II

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.75	1.00			
TGD	0.20	0.17	1.00		
piDAPH3	0.57	0.54	0.31	1.00	
SP	0.63	0.78	0.23	0.56	1.00

Región III

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.89	1.00			
TGD	0.93	0.87	1.00		
piDAPH3	0.90	0.85	0.93	1.00	
SP	0.80	0.78	0.80	0.80	1.00

Región IV

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.36	1.00			
TGD	0.24	0.19	1.00		
piDAPH3	0.14	0.17	0.13	1.00	
SP	0.15	0.25	0.07	0.11	1.00

G. intestinalis (división similitud en actividad 0.75)

Región I

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.76	1.00			
TGD	0.76	0.75	1.00		
piDAPH3	0.76	0.78	0.86	1.00	
SP	0.62	0.75	0.65	0.64	1.00

Región II

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.54	1.00			
TGD	0.44	0.49	1.00		
piDAPH3	0.38	0.48	0.54	1.00	
SP	0.45	0.65	0.44	0.36	1.00

Región III

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.84	1.00			
TGD	0.84	0.83	1.00		
piDAPH3	0.86	0.86	0.92	1.00	
SP	0.76	0.79	0.74	0.75	1.00

Región IV

	Radial	MACCS	TGD	piDAPH3	SP
Radial	1.00				
MACCS	0.43	1.00			
TGD	0.29	0.33	1.00		
piDAPH3	0.26	0.33	0.39	1.00	
SP	0.37	0.46	0.27	0.21	1.00

Tabla 9. Ejemplos de pares consenso en los maps SAS

Par	Similitud en actividad		Similitud estructural					SP ^c
	<i>T. vaginalis</i>	<i>G. intestinalis</i>	Radial	MACCS	TGD	piDAPH3	Promedio (σ)	
4_28^a	0.911	0.180	0.138	0.612	0.848	0.561	0.540 (0.295)	0.443
8_28^a	0.902	0.910	0.082	0.413	0.594	0.450	0.385 (0.216)	0.291
10_28^a	0.868	0.658	0.080	0.413	0.618	0.487	0.400 (0.229)	0.348
21_28^a	0.952	0.692	0.188	0.617	0.805	0.620	0.557 (0.262)	0.503
28_32^{a,b}	0.969	0.887	0.190	0.698	0.945	0.693	0.631 (0.317)	0.603
7_12^b	0.993	0.342	0.079	0.632	0.544	0.266	0.380 (0.254)	0.369
16_20^{a,b}	0.964	0.970	0.375	1.000	1.000	1.000	0.844 (0.313)	1.000
16_24^{a,b}	0.954	0.910	0.346	1.000	1.000	0.928	0.818 (0.317)	0.900
16_28^{a,b}	0.844	0.940	0.351	0.905	0.949	0.941	0.786 (0.291)	0.900
15_16^{a,b}	0.739	0.970	0.351	0.708	0.983	0.836	0.719 (0.270)	0.616
16_19^{a,b}	0.751	0.733	0.253	0.708	0.983	0.836	0.695 (0.315)	0.616
16_30^{a,b}	0.799	0.992	0.269	0.791	0.931	0.836	0.707 (0.297)	0.615
16_31^{a,b}	0.741	0.944	0.375	0.791	0.931	0.836	0.733 (0.246)	0.615
8_9^a	0.511	0.876	0.250	1.000	1.000	0.723	0.743 (0.354)	0.664
1_9^a	0.233	0.895	0.222	0.852	0.918	0.476	0.617 (0.327)	0.327
9_10^a	0.477	0.624	0.148	0.929	0.975	0.737	0.697 (0.380)	0.591
9_12^a	0.113	0.853	0.313	1.000	0.975	0.944	0.808 (0.331)	0.819
10_12^b	0.635	0.477	0.182	0.929	1.000	0.756	0.717 (0.371)	0.511

^aPar activo para *T. vaginalis*

^bPar activo para *G. intestinalis*

^cSimilitud en propiedades

Para *T. vaginalis*, muchos de los pares consenso con baja similitud estructural y alta similitud en actividad (región I) contienen el compuesto activo **28** ($CI_{50} = 86$ nM) como **4_28**, **8_28**, **10_28**, **21_28** y **32_28** (Tabla 9). La Figura 19A muestra la comparación de las estructuras químicas para pares selectos en la región I, además se muestran las actividades y medidas de similitud. En esta figura, los óvalos concéntricos indican diferentes grados de similitud estructural con respecto a la referencia **28**. Por ejemplo, los compuestos **8** y **10** son menos similares a **28** cuando se comparan con los compuestos **4**, **21** y **32**. Es interesante notar que **8** y **10** son menos activos que **4**, **21** y **32**. Dado a que todos los compuestos en la base de datos tienen el mismo núcleo, los pares en la región I pueden considerarse ejemplos de moléculas en las que cambios grandes en las cadenas laterales dirigen a compuestos con actividad similar (*side chain hopping*).

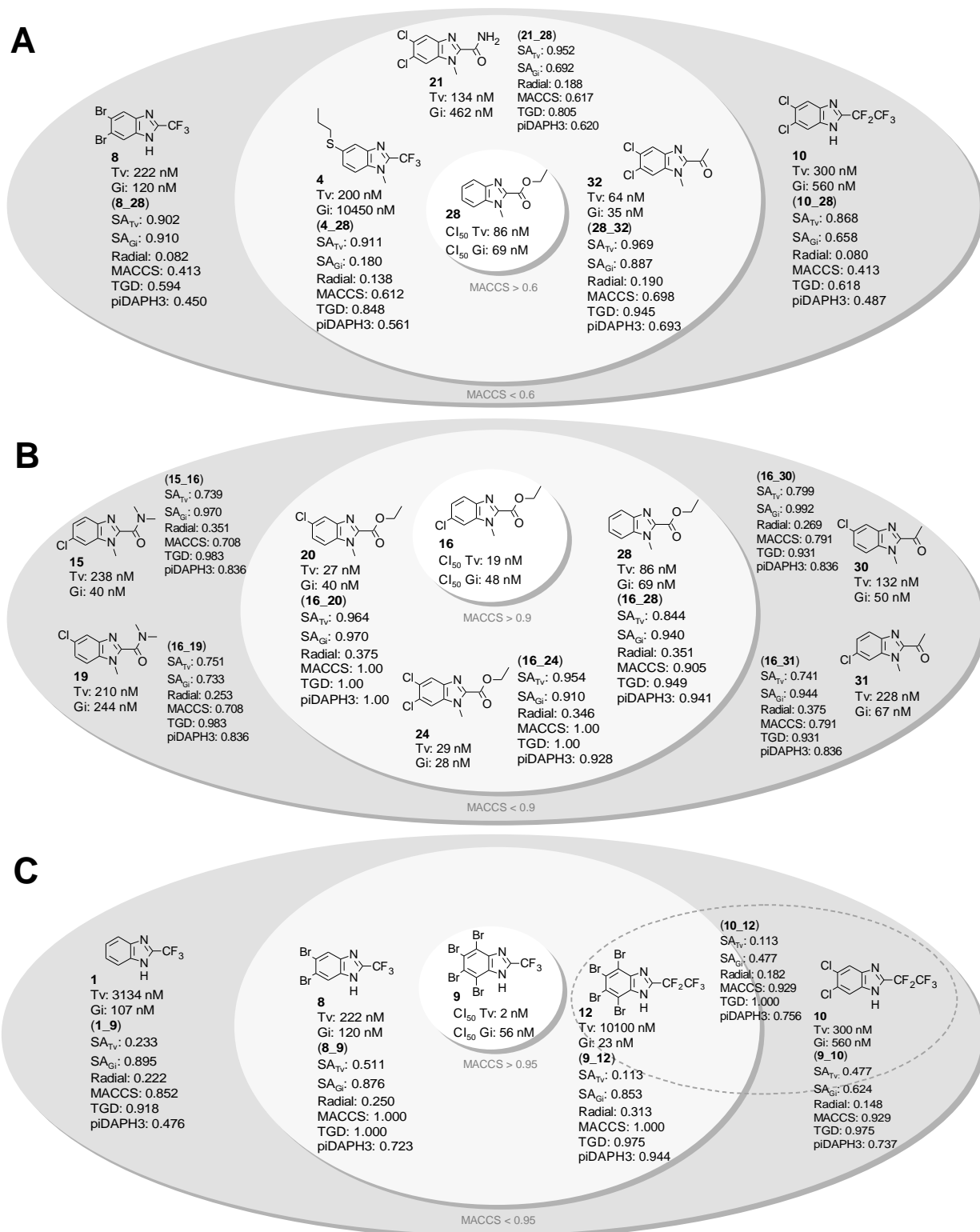


Figura 19. Compuestos incluidos en algunos pares consenso representativos para *T. vaginalis* y *G. intestinalis*. Los compuestos en los óvalos internos son más similares estructuralmente que los contenidos en los externos respecto a la molécula de referencia en el ovalo central. (A) Región I; (B) región II; (C) región IV (*activity cliffs*).

Algunos de los pares en la región I para *T. vaginalis* se encuentran en la misma región para *G. intestinalis*. Por ejemplo **8_28** y **28_32** con valores de similitud en actividad de 0.910 y 0.887, respectivamente (Tabla 9 y Figura 19). Estos resultados sugieren que las relaciones estructura-actividad son similares para ambos protozoarios. Los pares **4_28** y **7_12**, los cuales tienen baja similitud en actividad solamente contra *G. intestinalis* (0.180 y 0.342, respectivamente), indican que en algunos casos los mismos cambios en la estructura del bencimidazol producen diferentes efectos en la actividad contra *T. vaginalis* y *G. intestinalis*.

También se identificaron varios pares consenso en la región II para *T. vaginalis*; varios de esos pares contienen el compuesto activo **16** ($CI_{50} = 19$ nM). Algunos ejemplos son los pares **16_20**, **16_24** y **16_28** (Figura 19B). Es interesante que varias huellas digitales como MACCS, TGD y piDAPH3 son incapaces de distinguir entre isómeros posicionales como en el caso del par **16_20** (similitud = 1.0). Sin embargo, la huella digital radial puede diferenciar este par demostrando la alta resolución de esta representación [76]. Los pares donde ambos compuestos contienen éster etílico en posición 2 del bencimidazol (**16_20**, **16_24** y **16_28**) se ubican en la región lisa (región II); para estos pares, cambios en el patrón de sustitución con cloro en las posiciones 5 y 6 del núcleo de bencimidazol producen cambios pequeños en la actividad (similitud en actividad entre 0.844 y 0.964).

Los pares **16_15**, **16_19**, **16_30** y **16_31** están ubicados en la región II para *T. vaginalis* (Figura 19B). La similitud estructural de **15**, **19**, **30** y **31** decrece con respecto a **16** (0.739–0.799) en comparación con **20**, **24** y **28** (0.844–0.964). Estos resultados están de acuerdo con el principio de similitud e ilustran el SAR continuo asociado con los compuestos en la región II. Adicionalmente, la Figura 19B ilustra la importancia del sustituyente éster etílico en posición 2 del bencimidazol en la actividad contra *T. vaginalis*.

Muchos pares con alta similitud estructural y en actividad para *T. vaginalis* están también ubicados en la región II para *G. intestinalis* como se ilustra con los pares que contienen el compuesto **16** ($CI_{50} = 48$ nM, Tabla 9). Estos resultados indican que la sustitución con éster en posición 2 del bencimidazol incrementa la actividad contra

ambos parásitos y la sustitución con cloro en posiciones 5 y 6 no afecta de manera importante.

7.1.1.8. *Activity cliffs profundos y superficiales consenso, y activity cliffs aparentes*

La importancia de los *activity cliffs* ha sido discutida de manera amplia en la literatura [52–53, 78]. El estudio de *activity cliffs* es valioso ya que proporciona información sobre características estructurales específicas para lograr cambios importantes en la actividad. Los *activity cliffs consenso* son aquellos que pueden detectarse en diferentes representaciones moleculares [55]. En la Figura 19C se muestran ejemplos de *activity cliffs consenso* (región IV).

Para *T. vaginalis* se encontró sólo un *activity cliff profundo consenso*, el par **9_12** (*cliff* para todas las representaciones con más de dos unidades logarítmicas en diferencia de potencia). La diferencia estructural entre **9** y **12** es un grupo CF₂ en posición 2 del núcleo de bencimidazol (cambio de CF₃ por CF₂CF₃). Este cambio en la estructura produce un decremento en la actividad contra *T. vaginalis* desde CI₅₀ = 2 nM (**9**) hasta 10100 nM (**12**). Esto indica que la actividad de al menos uno de estos dos compuestos, o ambos, podría ser no predecible en un estudio QSAR [52]. Es interesante señalar que el par **9_12** no representa un *activity cliff* para *G. intestinalis* (CI₅₀ = 56 nM y 23 nM, respectivamente). Un segundo *activity cliff consenso* para *T. vaginalis* es el par **1_9**, detectado sólo por las huellas digitales radial y MACCS. Este es un ejemplo de un *cliff aparente* que pueden ser definidos como *activity cliffs* detectados sólo por algunas representaciones moleculares [55]. Un ejemplo más de *activity cliff aparente* es el par **9_10** identificado por MACCS y TGD. Un caso en el borde entre *activity cliff profundo* y *superficial* para *T. vaginalis* es **8_9** (similitud en actividad 0.511) identificado como *cliff* por radial, MACCS y TGD (pero no por piDAPH3). Sin embargo, MACCS y TGD no pueden distinguir entre este par (similitud = 1).

Se identificaron algunos *activity cliffs consenso* para *G. intestinalis*. Por ejemplo, el par **10_12** (Tabla 9 y Figuras 17 y 19C) fue identificado como *activity cliff profundo* solamente por las representaciones MACCS y TGD (*activity cliff aparente*). Como se discutió anteriormente, cuando la división para la similitud en actividad se ajusta a

0.75, el número de *activity cliffs* aumenta. Un ejemplo de *activity cliff superficial* (también aparente) es el par **9_10** identificado por MACCS y TGD. Este par fue también identificado como un *activity cliff aparente* para *T. vaginalis* (Tabla 9).

La presencia de pares de compuestos en regiones continuas y discontinuas del panorama de actividad para *T. vaginalis* y *G. intestinalis* reveló un SAR heterogéneo para ambos parásitos. El panorama de actividad para *T. vaginalis* está caracterizado por la presencia de dos *activity cliffs profundos consenso* para todas las representaciones analizadas y muchos puntos en la región del SAR continuo. En contraste, para *G. intestinalis* no se encontraron *activity cliffs consenso* para todas las representaciones pero se encontró un mayor número de *activity cliffs superficiales* comparado con *T. vaginalis*.

7.1.1.9. Modelos consenso del panorama de actividad: mapas SAS consenso

Los resultados obtenidos del análisis de los mapas SAS dependen de cada representación usada para obtener la similitud estructural [55, 77]. Sin embargo, los pares consenso encontrados para varias representaciones sugieren que es posible derivar, al menos de manera aproximada, modelos consenso para el panorama de actividad. Para esto, en este trabajo se empleó la fusión de datos de similitud estructural para producir los mapas SAS consenso [79]. Para cada par de compuestos se calculó el promedio y la desviación estándar para los valores de las huellas digitales radial, MACCS, TGD y piDAPH3 (las cuatro huellas digitales ortogonales seleccionadas). En la Figura 20 se muestran los mapas SAS consenso para *T. vaginalis* y *G. intestinalis*. En estos mapas, los puntos están coloreados por escala continua respecto al compuesto más activo en el par; además se muestran los puntos en tamaño proporcional al valor de la desviación estándar.

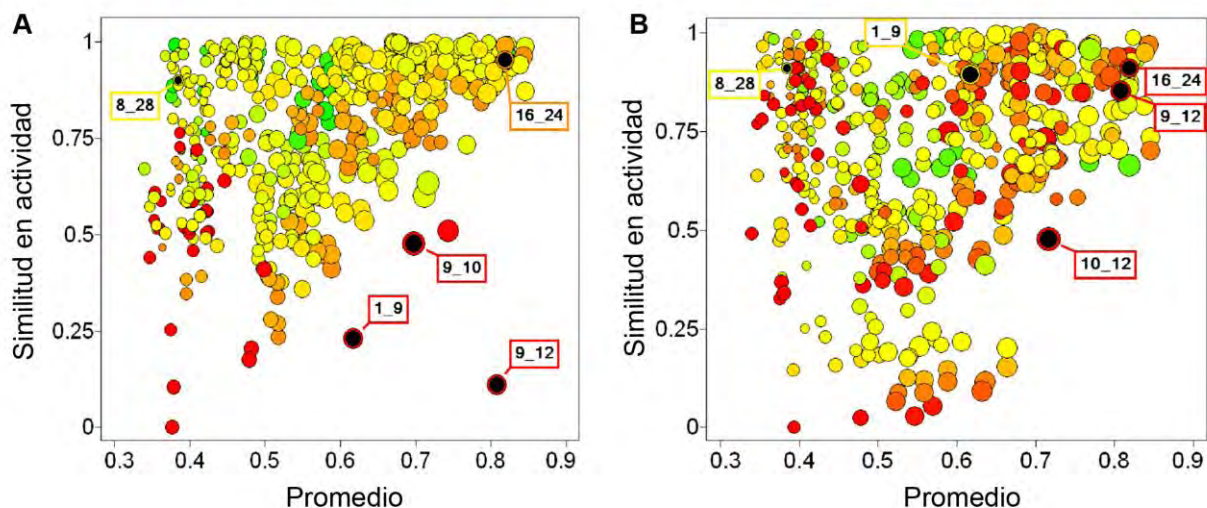


Figura 20. Mapas SAS consenso que muestran puntos con tamaño proporcional a la desviación estándar. (A) mapa para *T. vaginalis*; (B) mapa para *G. intestinalis*. Cada punto indica una comparación pareada para los 32 bencimidazoles (496 puntos). Los puntos están coloreados en escala continua que va desde verde (menos activo) hasta rojo (más activo). Los pares seleccionados están marcados en negro y etiquetados con la clave de los compuestos presentes.

En las Figuras 16E, 17E y 20 se muestra la posición relativa de algunos pares consenso previamente identificados por las huellas digitales radial, MACCS, TGD y piDAPH3 (Figuras 16A-D y 17A-D). Los valores de similitud promedio para las cuatro huellas digitales seleccionadas y la desviación estándar se muestra en la Tabla 9. En general, los pares en los mapas consenso ocupan una posición relativa similar (en las regiones I–IV) en comparación con los mapas obtenidos por las representaciones moleculares separadas (Figuras 16A-D y 17A-D). Por esta razón, se puede decir que los mapas SAS consenso capturan la información contenida en las representaciones independientes y proporcionan una buena aproximación a un modelo global del panorama de actividad de *T. vaginalis* y *G. intestinalis*. Basados en estos resultados, los mapas consenso pueden ser útiles para analizar bases de datos con otro tipo de actividades.

7.1.2. Panorama de actividad y selectividad para 55 derivados de bencimidazol

Tomando en cuenta los resultados obtenidos para 32 derivados de bencimidazol (1–32), se decidió expandir el estudio tomando en cuenta las estructuras (1–78) que incluyen 55 compuestos. A diferencia del grupo de moléculas usado para el

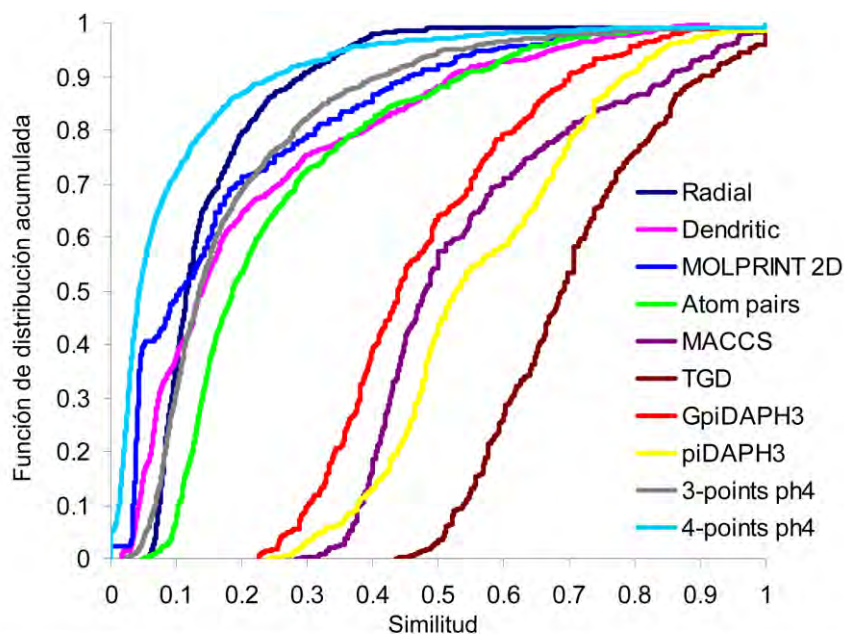
estudio anterior (1–32), el nuevo grupo de moléculas contiene compuestos con dos tautómeros (33–78). Se consideraron ambos tautómeros en el estudio dado a que se desconoce el grado en que cada uno de ellos contribuye en la actividad biológica. El nuevo grupo de moléculas contiene 14 compuestos con sustituyente alquiltio en posición 2, y nueve 2-(trifluorometil)bencimidazoles [16–19]. Con el fin de comparar el comportamiento de este nuevo grupo de moléculas, con el estudio anterior, se analizó de manera similar el panorama de actividad. Además se analizó el perfil de selectividad de este grupo de moléculas mediante los mapas de deferencia de actividad dual (DAD) propuestos por primera vez en este trabajo.

7.1.2.1. Selección de huellas digitales

Se consideraron inicialmente diez huellas digitales teniendo en cuenta los resultados del estudio anterior: radial, dendritic, MOLPRINT 2D, atom pairs, MACCS (166-bits), TGD, GpiDAPH3, piDAPH3, 3-, y 4-points ph4. En la Figura 21 se muestra la distribución para el nuevo grupo de moléculas y en la Tabla 10 se muestra la correlación entre las diez huellas digitales. Los resultados muestran que el nuevo grupo de 55 moléculas se comporta de una manera muy semejante al grupo de 32 moléculas. Basados en la distribución de los pares y la baja correlación (<0.75) se eligieron las representaciones dendritic, MACCS y piDAPH3. Es importante mencionar que la huella digital dendritic puede diferenciar entre formas tautoméricas.

Tabla 10. Matriz de correlación para las huellas digitales

	Radial	Dendritic	MOLPRINT 2D	Atom pairs	MACCS	TGD	GpiDAPH3	piDAPH3	3-points ph4	4-points ph4
Radial	1.00									
Dendritic	0.72	1.00								
MOLPRINT 2D	0.87	0.80	1.00							
Atom pairs	0.87	0.82	0.93	1.00						
MACCS	0.72	0.74	0.78	0.88	1.00					
TGD	0.56	0.67	0.58	0.70	0.69	1.00				
GpiDAPH3	0.59	0.74	0.66	0.72	0.65	0.68	1.00			
piDAPH3	0.55	0.69	0.61	0.65	0.56	0.64	0.91	1.00		
3-points ph4	0.63	0.77	0.70	0.75	0.71	0.68	0.78	0.75	1.00	
4-points ph4	0.60	0.71	0.65	0.70	0.65	0.60	0.72	0.69	0.96	1.00



Representación	Max	Q3	Mediana	Q1	Min	Promedio	σ
Radial	1.00	0.18	0.12	0.09	0.05	0.16	0.12
Dendritic	0.91	0.30	0.14	0.07	0.02	0.22	0.20
MOLPRINT 2D	1.00	0.25	0.11	0.04	0.00	0.18	0.19
Atom pairs	1.00	0.33	0.19	0.13	0.05	0.26	0.18
MACCS	1.00	0.64	0.48	0.42	0.28	0.55	0.18
TGD	1.00	0.80	0.69	0.59	0.44	0.70	0.14
GpiDAPH3	1.00	0.58	0.44	0.36	0.23	0.48	0.16
piDAPH3	1.00	0.69	0.53	0.46	0.24	0.57	0.16
3-points ph4	1.00	0.24	0.14	0.09	0.02	0.20	0.17
4-points ph4	1.00	0.12	0.04	0.02	0.00	0.10	0.15

Figura 21. Funciones de distribución acumulada para distintas representaciones de similitud estructural en 2D y 3D. La tabla resume la información de las distribuciones. Q3 y Q1 representan el tercer y primer cuartil respectivamente; σ representa la desviación estándar.

7.1.2.2. Panorama de actividad para el nuevo grupo de 55 moléculas

En la Figura 22 se muestran tres mapas SAS para cada protozooario construidos con las tres huellas digitales seleccionadas. Para cada protozooario se identificaron varios pares situados en la misma región para las tres representaciones (pares consenso). Ejemplos de pares consenso representativos se muestran marcados en negro en los mapas SAS.

Se identificaron también pares de compuestos en la misma región entre dos parásitos para las tres representaciones moleculares (pares consenso de doble blanco). La Tabla 11 resume el número de pares en cada región de los mapas SAS para cada protozooario y para doble blanco con distintas representaciones moleculares.

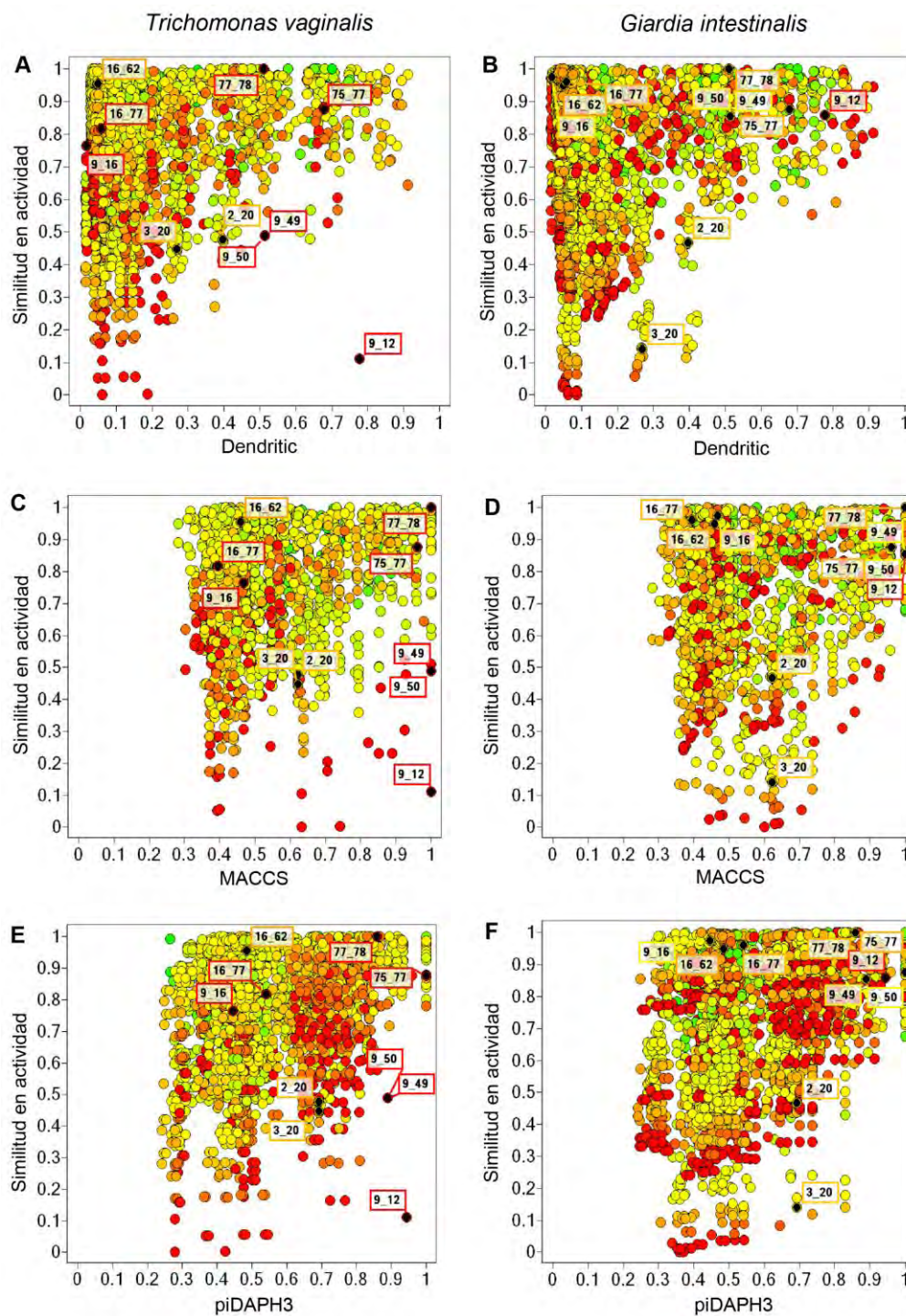


Figura 22. Mapas SAS para *T. vaginalis* y *G. intestinalis* con tres diferentes representaciones moleculares: (A) y (B) Dendritic; (C) y (D) MACCS; (E) y (F) piDAPH3. Cada punto significa una comparación pareada de los 55 bencimidazoles estudiados (78 estructuras, 3,003 puntos). Los puntos están codificados en escala continua de color donde los pares más activos se observan en rojo y los menos activos en verde. Los pares seleccionados están marcados en negro y etiquetados con el número correspondiente a los compuestos presentes.

Tabla 11. Distribución de puntos en las diferentes regiones de los mapas SAS^a

Representación	MSA ^b	I		II		III		IV	
		Pares Totales ^c	Pares activos ^d	Pares Totales	Pares activos	Pares Totales	Pares activos	Pares Totales	Pares activos
<i>T. vaginalis</i>									
Dendritic	0.26	1815	621	868	287	290	192	30	12
MACCS	0.54	1598	571	1085	337	237	158	83	46
piDAPH3	0.69	1959	635	724	237	289	177	31	27
Pares consenso ^e		1388	499	509	194	217	142	9	9
<i>G. intestinalis</i>									
Dendritic	0.24	1575	696	898	397	472	341	58	29
MACCS	0.55	1535	715	938	378	393	280	137	90
piDAPH3	0.67	1645	635	828	458	477	327	53	43
Pares consenso		1244	556	548	294	354	246	17	9
Pares consenso de doble blanco ^f		1104	300	491	160	122	70	4	4

^aLas regiones I–IV fueron definidas por la mediana de la similitud de los activos y un límite de 0.5 en similitud de actividad.

^bMediana de la similitud de los compuestos más activos ($pCI_{50} < 7.3$) para *T. vaginalis* y *G. intestinalis*.

^cNúmero total de puntos en la región correspondiente.

^dNúmero de puntos con al menos un compuesto activo en el par ($pCI_{50} < 7.3$).

^ePares consenso para cada protozoario usando tres diferentes representaciones.

^fPares consenso para ambos protozoarios usando tres diferentes representaciones.

Entre los resultados más importantes se encontró un alto número de pares consenso de doble blanco, en particular, en la región I y II (1104 y 491, respectivamente). Este resultado sugiere que muchos de los benzimidazoles estudiados pueden tener un mecanismo de acción común contra ambos protozoarios. También se identificaron cuatro *activity cliffs consenso* comunes para ambos parásitos (**2_16**, **2_20**, **3_16**, y **3_20**). De manera global, se encontró un SAR heterogéneo para ambos protozoarios. La Figura 23 muestra un ejemplo de mapa SAS para *T. vaginalis* con el nuevo grupo de moléculas donde se ilustran las estructuras de algunos pares consenso con sus valores de CI_{50} .

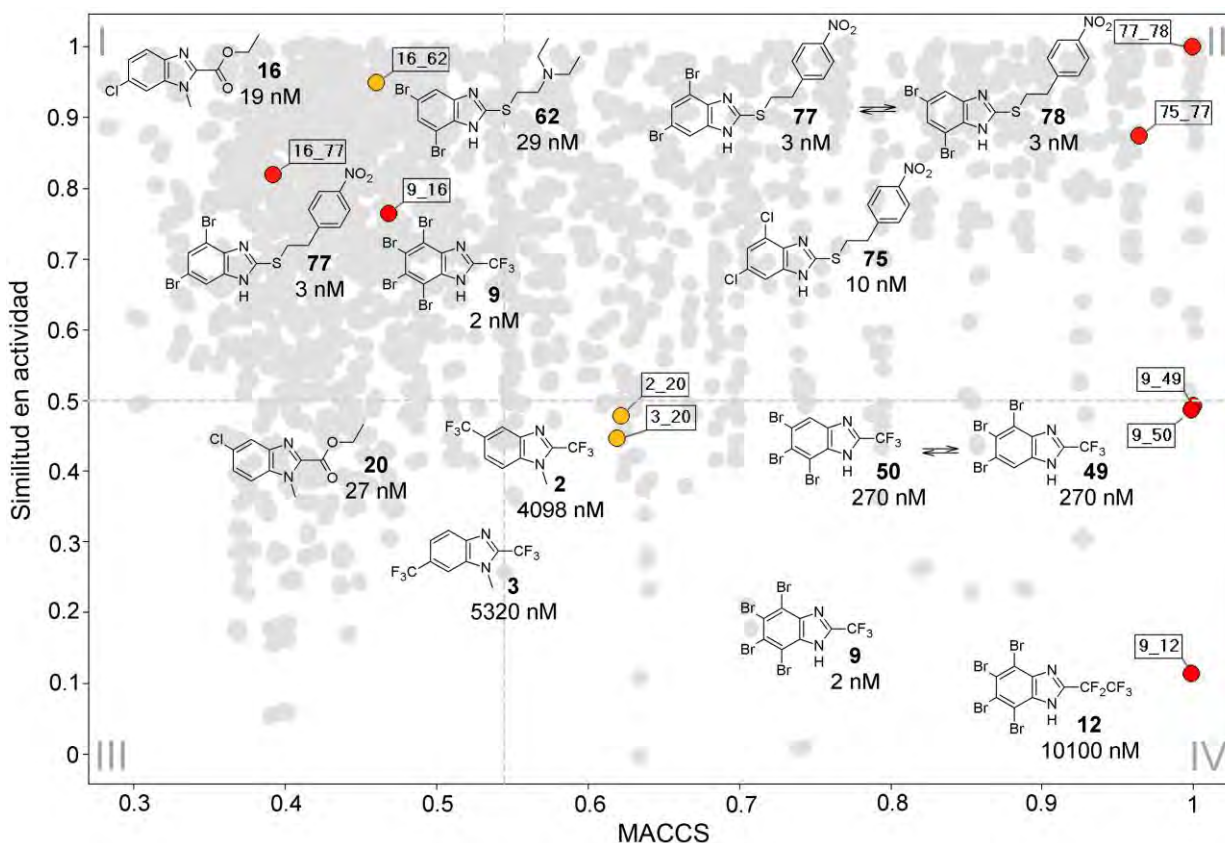


Figura 23. Mapa SAS representativo para *T. vaginalis* obtenido con la huella digital MACCS. La figura también muestra las estructuras químicas y la actividad tricomonocida en CI₅₀ de los pares consenso seleccionados en las tres regiones más importantes. Los pares seleccionados conservan el mismo código de color por actividad que en la Figura 22.

Los pares **9_16**, **16_62**, y **16_77**, en la región I, muestran ejemplos donde a pesar de realizar cambios grandes en las cadenas laterales o sustituyentes sólo se producen cambios pequeños en la actividad (*side chain hopping*). En la región II, el par **75_77** constituye un ejemplo de SAR continuo donde moléculas similares tienen actividades similares, y el par **77_78** representa a las dos formas tautoméricas donde la estructura es muy similar y la actividad se ha aproximado a ser la misma. El par **9_12**, en la región IV ilustra un *activity cliff* que también fue encontrado en el modelo construido con 32 compuestos. Ejemplos adicionales de *activity cliffs* mostrados en la Figura 23 son **2_20**, **3_20**, **9_49** y **9_50**. Aunque en estos últimos pares los cambios en la actividad son menos dramáticos que en **9_12**, aún se puede observar que cambios discretos en la estructura que conducen a cambios considerables en la actividad. Todos los pares discutidos en la Figura 23 tienen la misma posición relativa

en las representaciones dendríticas y piDAPH3. Además, la mayoría de esos pares ocupan la misma posición en el perfil de *G. intestinalis* con excepción de **9_12**, **9_49** y **9_50**.

7.1.2.3. Relación entre la actividad contra ambos parásitos y mapas de Diferencia de Actividad Dual (DAD)

En esta sección se comparó de manera sistemática la relación entre la actividad contra ambos parásitos. Los valores de diferencia de actividad entre dos parásitos (A y B) para cada compuesto se ilustran en la Tabla 2 ($\Delta pCl_{50} = pCl_{50} T. vaginalis (A) - pCl_{50} G. intestinalis (B)$). Muchos compuestos muestran actividad similar entre *T. vaginalis* y *G. intestinalis* (ΔpCl_{50} baja). Se observó que la mayoría de los compuestos más activos contra *T. vaginalis* ($pCl_{50} > 7$, $Cl_{50} < 100$ nM) también son muy activos contra *G. intestinalis*. En general, los compuestos más activos contra ambos protozoarios tienen sustituyentes grandes como alquiltio y carbetoxi en posición 2 (**16**, **20**, **23**, **24**, **28**, **32**, **55**, **59–68**, **71–78**). Además, para ambos parásitos, la sustitución con trifluorometilo en posición 2 da compuestos con actividades que van de pobres a moderadas ($pCl_{50} = 4.53–6.97$) con la excepción del compuesto **9** que tiene una actividad alta. Es interesante notar que los 1-metilbencimidazoles **13–32**, en casi todos los casos, tienen diferencia baja de actividad entre ambos protozoarios ($|\Delta pCl_{50}| = 0–0.77$, mediana $|\Delta pCl_{50}| = 0.205$). Sin embargo, también se identificaron moléculas con diferencia alta de actividad; entre ellas se encuentran nueve compuestos con diferencia de actividad mayor a la unidad logarítmica, de los cuales cinco compuestos tienen diferencia mayor de 1.4 unidades. La mayoría de los compuestos con la diferencia más grande tienen sustituidas totalmente las posiciones 5–6 con bromo o sustituyentes largos en 5 ó 6.

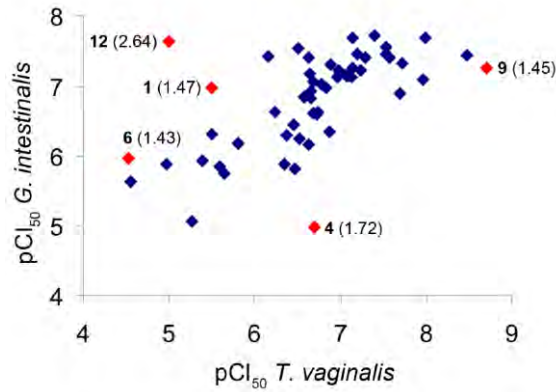


Figura 24. Correlación entre los valores de pCI_{50} para *T. vaginalis* y *G. intestinalis*. Los cinco compuestos con mayor diferencia de actividad están marcados en rojo. El valor absoluto de las diferencias $|\Delta pCI_{50}|$ está entre paréntesis.

La correlación para los valores de pCI_{50} entre ambos protozoarios se muestra en la Figura 24. De esta correlación se pueden identificar los cinco compuestos con mayor diferencia de actividad entre ambos protozoarios. Sin embargo, con sólo la correlación no es fácil asociar los cambios en actividad con cambios en estructura para ambos parásitos; con este fin se desarrollaron los mapas de diferencia de actividad dual (DAD). La forma general de los mapas DAD se muestra en la Figura 25. Estos mapas están basados en diferencias pareadas de actividad (calculadas usando la ecuación 8, $\Delta pCI_{50}(P)_{i,j}$); las diferencias de actividad contra un protozoario se grafican en el eje Y y las diferencias de actividad contra otro protozoario se grafican en el eje X.

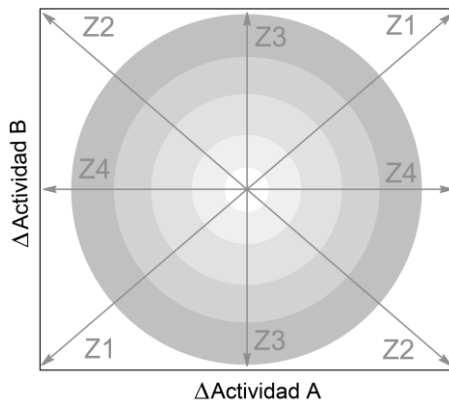


Figura 25. Forma general de un mapa DAD donde se muestran las cuatro zonas principales Z1–Z4. Los anillos concéntricos muestran diferentes magnitudes en diferencias de actividad; los anillos externos denotan mayor diferencia. En los pares cercanos a Z1 la actividad aumenta o decrece en proporción similar ante ambos parásitos (A y B). En pares cercanos a Z2 los cambios en actividad de A son inversamente proporcionales a los cambios en B. En los pares cercanos a Z3 ó Z4 se observan pares donde la actividad se mantiene constante para un protozoario (A o B) mientras que la actividad cambia contra el otro (A o B). Pares localizados en la parte central del mapa indican que no hay cambios en actividad contra A y B.

Los mapas DAD proveen una caracterización visual del perfil de selectividad para la serie de compuestos evaluada contra dos parásitos, A y B (pueden aplicarse para otros sistemas). En general, en la Figura 25 se pueden distinguir cuatro zonas Z1–Z4. Los anillos concéntricos en el diagrama general denotan diferentes magnitudes en diferencia de actividad; los anillos externos representan mayor diferencia. Los pares de compuestos sobre o cercanos a Z1, indican que la actividad de los dos compuestos en el par aumenta o disminuye en proporciones similares contra A y B. Por esto, los pares cercanos a la zona Z1 tienen SAR similar contra ambos parásitos. Pares sobre o cercanos a Z2, indican que los cambios en actividad son opuestos contra A y B. Los pares sobre Z2 son asociados con un SAR inverso; en otras palabras, cuando la actividad aumenta para un parásito disminuye para el otro con la misma modificación estructural. Los pares cercanos a Z3 y Z4, denotan pares de moléculas sin cambios en actividad contra un blanco (A o B), pero con cambios contra el segundo blanco (A o B, respectivamente). Los datos localizados en la intersección de Z1–Z4, o cercanos a ésta, indican que no hay cambios en actividad para A o B con esa modificación estructural.

En la Figura 26 se muestra un ejemplo de mapa DAD para derivados de bencimidazol evaluados contra *T. vaginalis* y *G. intestinalis*.

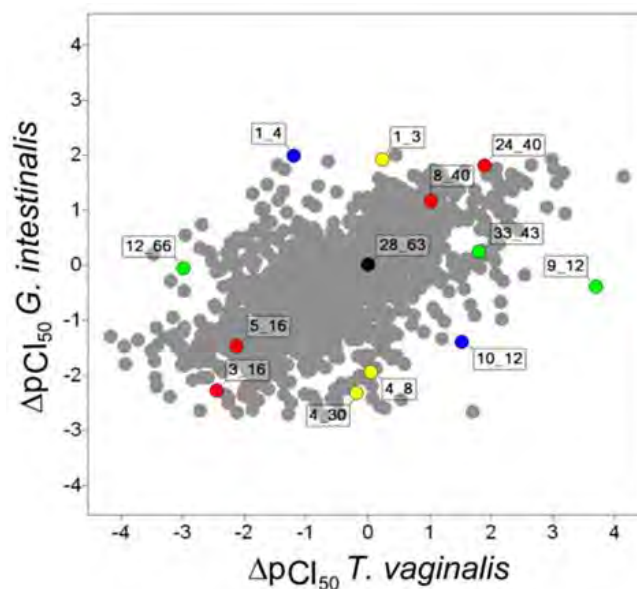


Figura 26. Mapa DAD para la base de datos de derivados de bencimidazol con actividad contra *T. vaginalis* y *G. intestinalis*. Los puntos en color significan ejemplos de pares cercanos a las diferentes zonas: Z1, rojo; Z2, azul; Z3, amarillo; Z4, verde; e intersección Z1–Z4, negro.

La gráfica contiene 3,003 puntos que representan todas las comparaciones pareadas (se omitieron duplicados). Los pares selectos están coloreados según la zona a la que se encuentran cercanos: Los puntos rojos están cercanos a Z1, los azules a Z2, los amarillos a Z3, y los verdes a Z4. El punto negro está aproximadamente en el centro del mapa y cercano a la intersección entre las cuatro zonas.

La Figura 27 muestra las estructuras químicas de los pares selectos y sus valores de CI_{50} (nM) contra *Trichomonas vaginalis* y *Giardia intestinalis*.

El par **28_63**, localizado cerca del centro del gráfico, ejemplifica un par de moléculas donde cambios estructurales no modifican significativamente la actividad para *T. vaginalis* (CI_{50} = 86 contra 85 nM) o para *G. intestinalis* (CI_{50} = 69 contra 72 nM). Algo interesante es que el par **28_63** es también un ejemplo de *side chain hopping* debido a que el patrón de sustitución es diferente en ambas moléculas y la actividad contra ambos parásitos es muy similar.

Los pares **3_16**, **5_16**, **8_40**, y **24_40** son ejemplos de puntos cercanos a Z1. Los cambios estructurales en estos pares están asociados a cambios similares en actividad contra ambos protozoarios. Por ejemplo, para el par **3_16** el cambio de trifluorometilo (**3**) por carbetoxi (**16**) en posición 2 y de trifluorometilo (**3**) por cloro (**16**) en posición 6, está asociado con el aumento de actividad contra ambos parásitos. Se pueden obtener conclusiones similares con otros pares cercanos a Z1. Por otra parte, es importante notar que la diferencia de actividad para ambos parásitos es más grande en los pares **3_16** y **24_40** (más alejados del centro del gráfico) comparados con los pares **5_16** y **8_40** (localizados más cerca del centro del gráfico).

Los pares **1_4** y **10_12** son ejemplos de pares cercanos a la línea Z2. El cambio estructural en cada par está asociado con un cambio inverso en la actividad. Por ejemplo, comparando el par **1_4**, la *N*-metilación y la sustitución en posición 5 con un grupo propiltio (**4**) están asociadas con un aumento significativo en la actividad para *T. vaginalis*. Sin embargo, los mismos cambios estructurales están asociados con un decremento significativo para *G. intestinalis*. Nótese que **1** es selectivo para *G. intestinalis* y **4** es selectivo para *T. vaginalis*.

Los puntos 1_3, 4_8, y 4_30 son ejemplos de pares de compuestos cercanos a Z3. En estos puntos, la actividad contra *T. vaginalis* es similar mientras que la actividad contra *G. intestinalis* cambia. En contraste, en los pares 9_12, 12_66, y 33_43, cercanos a Z4, indican pares con actividad similar contra *G. intestinalis* pero con cambios para *T. vaginalis*. Uno de los ejemplos más importantes es el par 9_12, con actividad similar contra *G. intestinalis* pero con un cambio de 5050 veces en la CI_{50} para *T. vaginalis*.

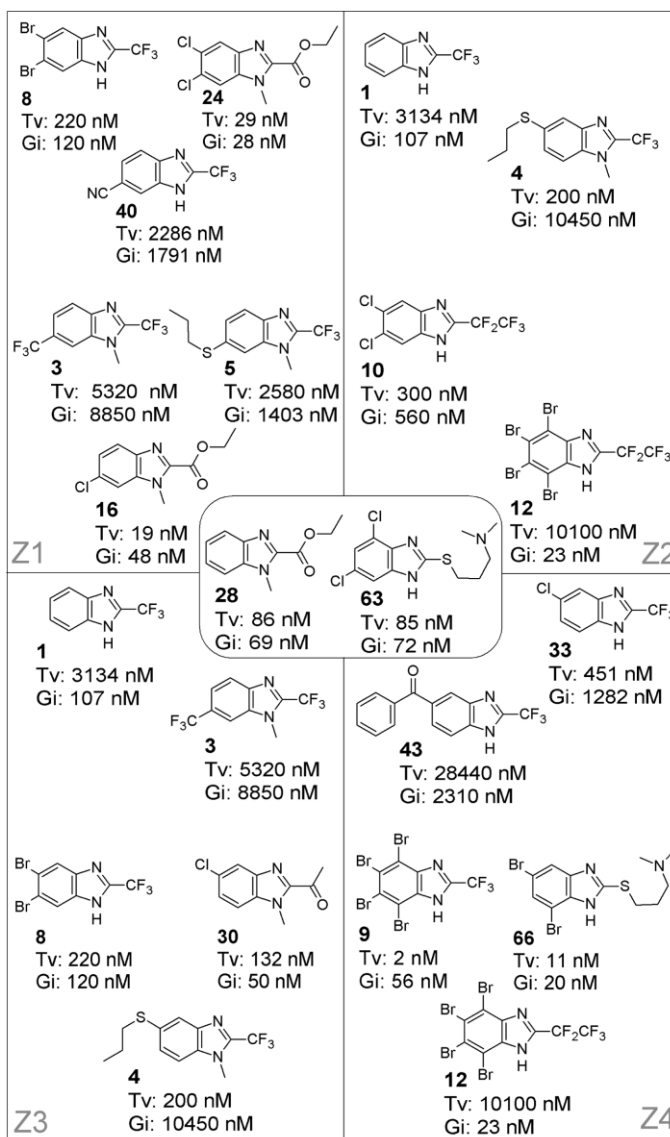


Figura 27. Estructuras químicas de los compuestos incluidos en los pares selectos con los valores de CI_{50} para *T. vaginalis* (Tv) y *Giardia intestinalis* (Gi).

7.1.2.4. Activity cliffs con mapas DAD

Debido a que los mapas DAD están contruidos por comparaciones pareadas, es posible añadir información de similitud estructural. La Figura 28 muestra un ejemplo de mapa DAD, para los derivados de bencimidazol, donde se agrega la similitud estructural usando la huella digital dendritic. Se usó una escala continua de color donde el verde significa compuestos menos similares y en rojo se señalan pares de compuestos con mayor similitud.

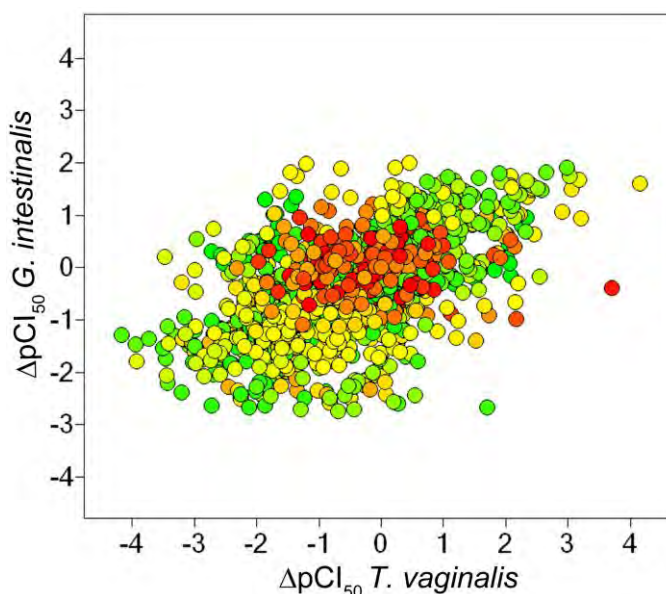


Figura 28. Mapa DAD con pares coloreados por similitud estructural. Los colores se encuentran en escala continua de verde (menos similar) a rojo (más similar).

En la Figura 29 se muestran en color 925 pares con alta similitud molecular (similitud más alta que la mediana de la similitud para los compuestos activos). Se pueden obtener mapas similares usando distintas huellas digitales como MACCS y piDAPH3. En general, se observa que compuestos con alta similitud estructural tienden a agruparse cerca del centro del gráfico. Esto quiere decir que muchos pares de bencimidazoles con alta similitud estructural tienen diferencias de actividad bajas con ambos parásitos. Este resultado concuerda con el principio de similitud (moléculas similares tienen actividades similares) y sugiere que varias moléculas tienen SAR continuo contra ambos parásitos como se observó en los mapas SAS. Sin embargo, también se detectaron pares de moléculas similares en estructura y con diferencias de

potencia grandes. Estos últimos pares brindan información importante para optimizar moléculas contra ambos parásitos o generar selectividad.

Los pares **3_16** y **4_30** en las regiones Z1 y Z3, respectivamente, tienen alta similitud molecular medida con la huella digital dendritic. El par **9_12** en la región Z4 también tiene una alta similitud molecular (punto rojo alejado del centro del mapa) y la diferencia de potencia contra *T. vaginalis* es muy amplia. Sin embargo, la diferencia de actividad contra *G. intestinalis* es pequeña. Por esto, es fácil deducir que el par **9_12** es un *activity cliff* para *T. vaginalis* y está en la región lisa para *G. intestinalis* en los mapas SAS.

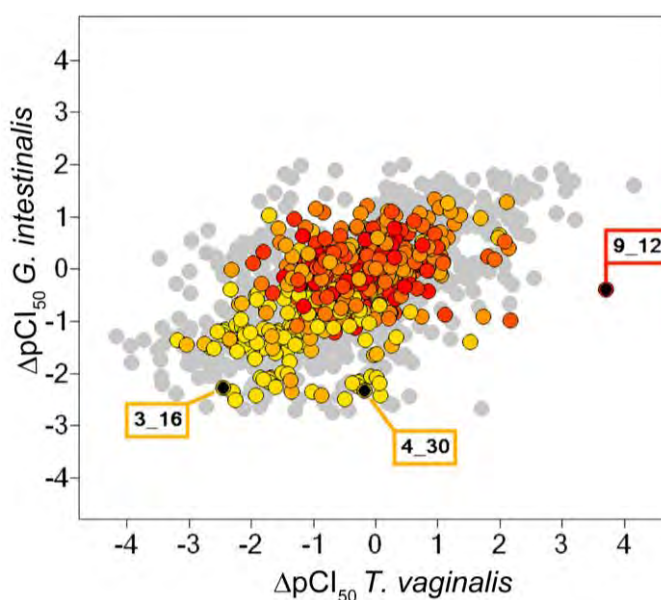


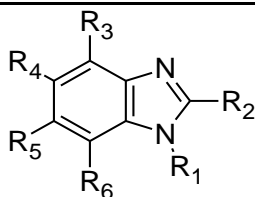
Figura 29. Mapa DAD que muestra 925 pares con alta similitud estructural. Los pares selectos están marcados en negro y etiquetados con el número de los compuestos que los componen.

Cabe mencionar que los mapas DAD son una nueva herramienta implementada por primera vez en este trabajo y se pueden aplicar para cualquier tipo de sistema facilitando el estudio de las relaciones estructura-actividad en bases de datos extensas.

7.2. Relaciones cuantitativas estructura–actividad de derivados de bencimidazol con actividad tricomonocida

En el presente trabajo se estudió una base de datos de 70 derivados de bencimidazol activos contra *T. vaginalis* usando las metodologías de CoMFA y CoMSIA (Tabla 12) [11, 16–21].

Tabla 12. Estructura y actividad tricomonocida de los derivados de bencimidazol usados en los estudios QSAR-3D



	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	pCl ₅₀ ^a
1	H	CF ₃	H	H	H	H	5.50
2	H	CF ₃	H	Cl	H	H	6.35
3 ^b	H	CF ₃	H	F	H	H	5.50
4	H	CF ₃	H	CF ₃	H	H	6.63
5	H	CF ₃	H	CN	H	H	5.64
6	CH ₃	CF ₃	H	CF ₃	H	H	5.39
7	CH ₃	CF ₃	H	H	CF ₃	H	5.27
8	H	CF ₃	H	SCH ₂ CH ₂ CH ₃	H	H	6.46
9	CH ₃	CF ₃	H	SCH ₂ CH ₂ CH ₃	H	H	6.70
10	CH ₃	CF ₃	H	H	SCH ₂ CH ₂ CH ₃	H	5.59
11 ^b	H	CF ₃	H	COC ₆ H ₅	H	H	4.55
12	CH ₃	CF ₃	H	COC ₆ H ₅	H	H	4.53
13	CH ₃	CF ₃	H	H	COC ₆ H ₅	H	4.97
14	H	H	H	H	H	H	6.44
15	H	CH ₃	H	H	H	H	6.52
16	H	NH ₂	H	H	H	H	6.54
17	H	SH	H	H	H	H	6.71
18	H	H	H	Cl	H	H	6.69
19	H	CH ₃	H	Cl	H	H	6.79
20	H	NH ₂	H	Cl	H	H	6.71
21	H	SH	H	Cl	H	H	6.87
22 ^b	H	SCH ₃	H	Cl	H	H	7.03
23	H	NH ₂	H	Cl	Cl	H	6.98
24	H	CF ₃	H	Br	H	H	5.80
25	H	CF ₃	H	Br	Br	H	6.66
26	H	CF ₃	Br	H	Br	H	6.72
27 ^b	H	CF ₃	Br	Br	Br	H	6.57
28	H	CF ₃	Br	Br	Br	Br	8.70
29	H	C ₂ F ₅	H	Cl	Cl	H	6.52
30	H	CF ₃	H	NO ₂	NO ₂	H	6.24
31	H	C ₂ F ₅	Br	Br	Br	Br	5.00
32	H	SCH ₂ CH ₂ OH	Cl	H	Cl	H	6.16
33	H	SCH ₂ CH ₂ OH	Br	H	Br	H	6.83

34	H	SCH ₂ CH ₂ N(CH ₃) ₂	Cl	H	Cl	H	7.14
35	H	SCH ₂ CH ₂ N(CH ₃) ₂	Br	H	Br	H	6.51
36 ^b	H	SCH ₂ CH ₂ N(C ₂ H ₅) ₂	Cl	H	Cl	H	7.39
37	H	SCH ₂ CH ₂ N(C ₂ H ₅) ₂	Br	H	Br	H	7.53
38	H	SCH ₂ CH ₂ CH ₂ N(CH ₃) ₂	Cl	H	Cl	H	7.07
39	H	SCH ₂ CH ₂ CH ₂ N(CH ₃) ₂	Br	H	Br	H	7.99
40	H	SCH ₂ CH ₂ -(<i>N</i> -piperidil)	Cl	H	Cl	H	7.14
41 ^b	H	SCH ₂ CH ₂ -(<i>N</i> -piperidil)	Br	H	Br	H	7.69
42	H	SCH ₂ CH ₂ -(<i>N</i> -morfolinil)	Cl	H	Cl	H	7.29
43	H	SCH ₂ CH ₂ -(<i>N</i> -morfolinil)	Br	H	Br	H	7.23
44	H	SCH ₂ CH ₂ -(<i>p</i> -nitrofenil)	Cl	H	Cl	H	7.96
45	H	SCH ₂ CH ₂ -(<i>p</i> -nitrofenil)	Br	H	Br	H	8.48
46 ^b	CH ₃	CONH ₂	H	H	Cl	H	6.96
47	CH ₃	CONHCH ₃	H	H	Cl	H	6.98
48 ^b	CH ₃	CON(CH ₃) ₂	H	H	Cl	H	6.63
49	CH ₃	COOCH ₂ CH ₃	H	H	Cl	H	7.72
50	CH ₃	CONH ₂	H	Cl	H	H	6.73
51	CH ₃	CONHCH ₃	H	Cl	H	H	6.45
52	CH ₃	CON(CH ₃) ₂	H	Cl	H	H	6.68
53 ^b	CH ₃	COOCH ₂ CH ₃	H	Cl	H	H	7.57
54 ^b	CH ₃	CONH ₂	H	Cl	Cl	H	6.87
55	CH ₃	CONHCH ₃	H	Cl	Cl	H	6.65
56	CH ₃	CON(CH ₃) ₂	H	Cl	Cl	H	7.12
57	CH ₃	COOCH ₂ CH ₃	H	Cl	Cl	H	7.53
58	CH ₃	CONH ₂	H	H	H	H	6.78
59 ^b	CH ₃	CONHCH ₃	H	H	H	H	6.98
60	CH ₃	CON(CH ₃) ₂	H	H	H	H	6.37
61	CH ₃	COOCH ₂ CH ₃	H	H	H	H	7.07
62	CH ₃	NHCOOCH ₃	H	H	H	H	5.88
63	CH ₃	NHCOOCH ₃	H	Cl	H	H	6.36
64	CH ₃	NHCOOCH ₃	H	H	Cl	H	6.46
65 ^b	CH ₃	NHCOOCH ₃	H	Cl	Cl	H	6.27
66	H	SCH ₃	H	CONH ₂	H	H	6.67
67	H	SCH ₃	H	CONHCH ₃	H	H	6.79
68	H	SCH ₃	H	CON(CH ₃) ₂	H	H	7.54
69 ^b	H	SCH ₃	H	CONHCH ₂ CH ₃	H	H	7.26
70	H	SCH ₃	H	CON(CH ₂ CH ₃) ₂	H	H	6.05

^aLos datos de actividad se expresan en $-\log Cl_{50} = pCl_{50}$.

^bCompuestos incluidos en el grupo de prueba.

Una etapa fundamental para el desarrollo exitoso de modelos QSAR-3D es la selección de confórmeros, especialmente cuando el mecanismo de acción o sitio blanco es desconocido y las moléculas en estudio tienen varios enlaces rotables [22]. No obstante, muchos modelos construidos con conformaciones de mínima energía han mostrado buenos resultados cuantitativos en varios trabajos previos [80–82]. Adicionalmente, es muy importante tomar en cuenta el equilibrio tautomérico ya que una parte importante de las moléculas consideradas en este estudio poseen dos o tres formas tautoméricas relevantes. Es conocido que diferentes tautómeros pueden tener

diferente carácter hidrófobo, pKa, forma tridimensional y propiedades electrostáticas. Además, en algunos casos, el tautómero predominante en solución no es el biológicamente activo, especialmente cuando las barreras de interconversión entre tautómeros son pequeñas [83–85]. De manera general, para seleccionar formas tautoméricas representativas, primero se construyó un modelo de CoMFA que contiene moléculas sin tautómeros. Este modelo se usó para estimar la actividad de las moléculas en equilibrio tautomérico y seleccionar aquella representación cuya actividad estimada sea la más cercana a la actividad experimental [48]. Las formas tautoméricas más importantes consideradas en este trabajo y su clasificación se muestran en la Figura 30.

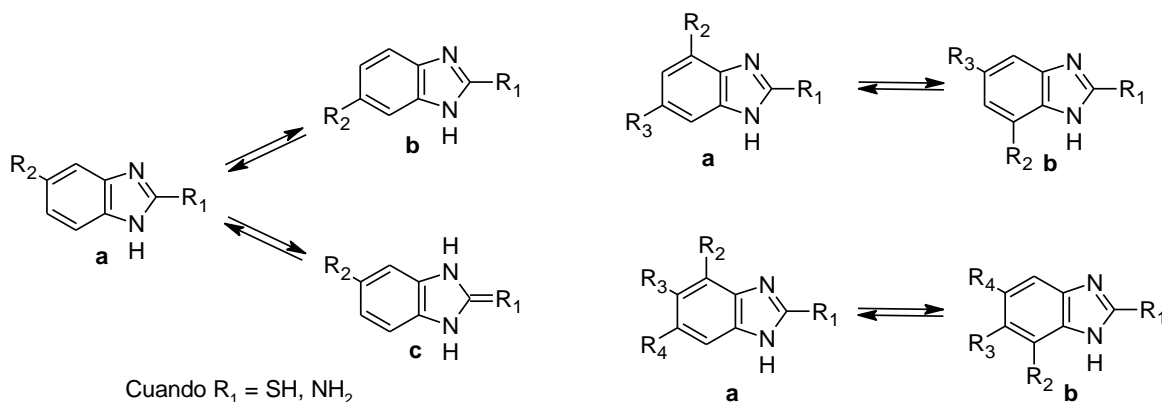


Figura 30. Formas tautoméricas consideradas este trabajo y su identificación basada en su patrón de sustitución.

7.2.1. Modelos construidos para selección de tautómeros

Como primer paso se construyó el modelo 1 incluyendo 28 moléculas sin tautómeros. Las moléculas se usaron en sus conformaciones de mínima energía y con las cargas derivadas del potencial electrostático. Los resultados para este modelo muestran baja correlación y poder predictivo basados en sus valores bajos de r^2 y q^2 , ($r^2 = 0.440$ y $q^2 = 0.043$; Tabla 13). Es interesante resaltar que se observó que los compuestos **28** y **31** son responsables de la pobre correlación.

Tabla 13. Resumen de resultados para los modelos de CoMFA 1–5

	Modelos CoMFA				
	1	2	3	4	5
q^2 ^a	0.043	0.503	0.465	0.472	0.433
PRESS ^b	0.895	0.615	0.638	0.618	0.640
N ^c	28	26	26	26	26
Noc ^d	1	6	6	5	5
r^2 ^e	0.443	0.975	0.975	0.964	0.971
S ^f	0.683	0.138	0.137	0.161	0.144
F ^g	20.670	123.045	124.608	107.995	135.622
Estérico ^h	0.518	0.501	0.509	0.520	0.555
Electrostático ⁱ	0.482	0.499	0.491	0.480	0.445

^aCoefficiente de determinación de la validación cruzada.

^bError estándar de las predicciones.

^cNúmero de compuestos en el grupo de trabajo.

^dNúmero óptimo de componentes principales.

^eCoefficiente de determinación no cruzado.

^fError estándar estimado.

^gPrueba de F.

^hContribución del parámetro estérico.

ⁱContribución del parámetro electrostático.

Posteriormente se construyó el modelo 2 con 26 compuestos (sin incluir **28** y **31**). Los resultados muestran un notable incremento en los valores de r^2 y q^2 ($r^2 = 0.975$ y $q^2 = 0.503$). Este modelo puede considerarse aceptable para usarse en la predicción de la actividad tricomonocida ($q^2 > 0.5$). Es interesante señalar que compuestos con comportamiento similar a **28** y **31** han sido ampliamente estudiados y han sido asociados con diferentes mecanismos de acción, distintos modos de unión, sitios de unión flexibles [86–88] y *activity cliffs* [52, 55]; en estos casos, las relaciones estructura-actividad no pueden explicarse por medio de los modelos QSAR convencionales. Cabe mencionar que los compuestos **28** y **31** fueron previamente detectados por formar *activity cliffs* en los mapas SAS (par **9_12** en la Figura 23). Aunque en este trabajo no existe evidencia experimental que explique el comportamiento de los compuestos **28** y **31**, son los únicos compuestos sustituidos totalmente en las posiciones 7–8 con átomos de bromo; por lo que esta característica puede ser un patrón importante para que estos compuestos presenten un SAR discontinuo.

Otra característica importante a tomar en cuenta es que los compuestos con sustituyentes grandes en las posiciones 5 y 6 (por ejemplo **12** y **13**) tienen dos

rotámetros de mínima energía. Los sustituyentes en estos rotámeros están orientados a lados opuestos del plano definido por el núcleo de bencimidazol. Por esta razón, además del modelo 2, se construyeron los modelos 3–5 agrupando las moléculas mediante la orientación de sus rotámeros como se muestra en la Figura 31.

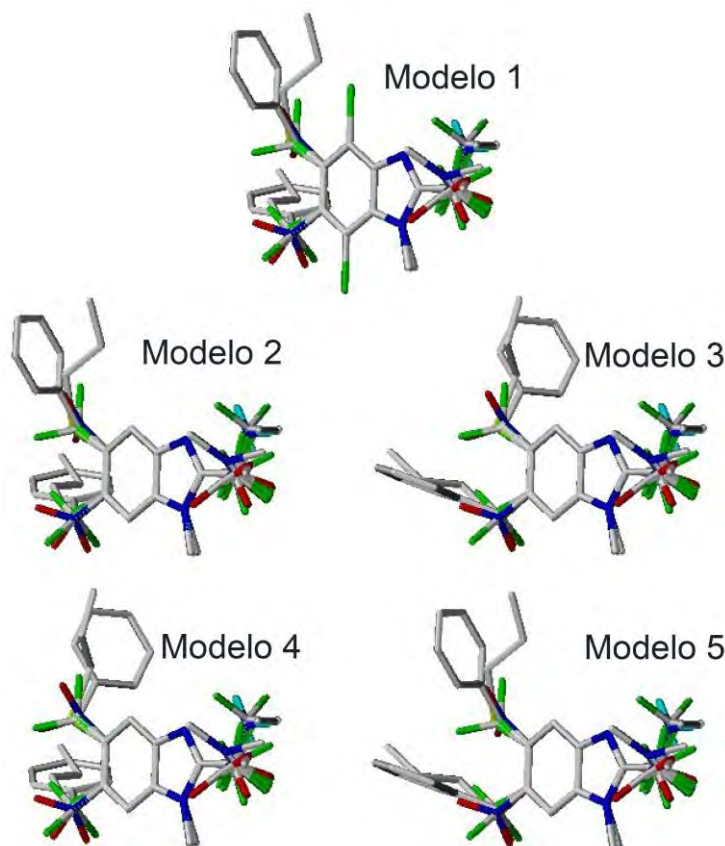


Figura 31. Orientación de los sustituyentes en las posiciones 5 y 6 para los modelos 1–5. Los hidrógenos no polares se removieron para mayor claridad.

Los mejores resultados se observaron conservando la orientación de los sustituyentes como se muestra en el modelo 2 de la Figura 31; siendo el único modelo que mostró $q^2 > 0.5$. Este modelo se usó posteriormente para estimar la actividad de cada representación en moléculas con dos o más formas tautoméricas y asignar el tautómero con mejor predicción, respecto a la actividad experimental, como responsable de la actividad. Los valores de actividad experimental y calculada para cada forma tautomérica se muestran en la Tabla 14.

Tabla 14. Valores de pCl₅₀ experimentales y calculados con modelo 2 para moléculas con tautómeros

	Experimental	Calculado	Residual		Experimental	Calculado	Residual
2a		5.72	0.63	33a*		7.07	-0.24
2b*	6.35	5.94	0.41	33b	6.83	6.56	0.27
3a*		5.63	-0.13	34a*		7.21	-0.07
3b	5.50	5.82	-0.32	34b	7.14	6.66	0.48
4a		5.60	1.03	35a		7.20	-0.69
4b*	6.63	5.61	1.02	35b*	6.51	6.70	-0.19
5a*		5.61	0.03	36a*		7.26	0.13
5b	5.64	5.92	-0.28	36b	7.39	6.83	0.56
8a		7.13	-0.67	37a*		7.21	0.32
8b*	6.46	6.03	0.43	37b	7.53	6.92	0.61
11a*		4.88	-0.33	38a*		7.19	-0.12
11b	4.55	5.35	-0.80	38b	7.07	6.92	0.15
16a*		6.90	-0.36	39a*		7.22	0.77
16c	6.54	5.69	0.85	39b	7.99	6.65	1.34
17a*		6.37	0.34	40a*		7.32	-0.18
17c	6.71	5.01	1.70	40b	7.14	6.74	0.40
18a*		6.69	0.00	41a*		7.30	0.39
18b	6.69	6.93	-0.24	41b	7.69	6.77	0.92
19a*		6.68	0.11	42a*		7.29	0.00
19b	6.79	6.92	-0.13	42b	7.29	6.75	0.54
20a*		7.06	-0.35	43a*		7.22	0.01
20b	6.71	7.30	-0.59	43b	7.23	6.78	0.45
20c		6.02	0.69	44a*		7.45	0.51
21a		6.56	0.31	44b	7.96	6.89	1.07
21b*	6.87	6.79	0.08	45a*		7.45	1.03
21c		5.20	1.67	45b	8.48	6.96	1.52
22a		6.90	0.13	66a*		6.65	0.02
22b*	7.03	7.15	-0.12	66b	6.67	7.06	-0.39
23a*		7.44	-0.46	67a*		6.92	-0.13
23c	6.98	6.26	0.72	67b	6.79	6.98	-0.19
24a*		6.18	-0.38	68a		6.63	0.91
24b	5.80	6.32	-0.52	68b*	7.54	6.93	0.61
26a*		6.56	0.16	69a*		6.71	0.55
26b	6.72	6.00	0.72	69b	7.26	6.59	0.67
27a*		6.44	0.13	70a*		6.28	-0.23
27b	6.57	6.02	0.55	70b		6.47	-0.42
32a		7.06	-0.90				
32b*	6.16	6.48	-0.32				

*Tautómero con el valor de actividad estimada más cercano al valor experimental.

7.2.2. Modelos de CoMFA y CoMSIA usando confórmers de mínima energía

Para generar los modelos de CoMFA y CoMSIA usando confórmers de mínima energía, se usaron 26 compuestos sin tautómeros y 29 compuestos con tautómeros, en su forma seleccionada mediante cálculos QSAR (Tabla 14), para dar un total de 55 moléculas.

Tabla 15. Resumen de resultados para los modelos de CoMFA 6–12

Cargas	Modelos CoMFA			Modelos CoMSIA			
	6	7	8	9	10	11	12
	Electrostáticas	Mulliken	Gasteiger-Huckel	Electrostáticas			
q^2	0.634	0.602	0.627	0.617	0.507	0.503	0.642
PRESS	0.468	0.488	0.472	0.474	0.527	0.535	0.454
N	55	55	55	55	55	55	55
Noc	5	5	5	4	2	3	3
r^2	0.936	0.927	0.928	0.915	0.573	0.706	0.858
S	0.196	0.209	0.208	0.223	0.491	0.412	0.286
F	142.893	124.974	125.819	134.483	34.844	40.730	102.338
Estérico	0.525	0.554	0.556	0.231			0.084
Electrostático	0.475	0.446	0.444	0.769			0.285
Donador ^j					0.596		0.227
Aceptor ^k					0.404		0.177
Hidrófobo ^l						1	0.226
R^{2m}	0.890	0.890	0.842	0.873	0.562	0.759	0.809
R_0^{2n}	0.828	0.843	0.809	0.777	0.509	0.663	0.749
k^o	1.002	0.996	1.004	0.993	0.994	0.999	0.998
$(R^2 - R_0^2)/R^2$	0.069	0.052	0.039	0.110	0.094	0.126	0.074
R_0^{2p}	0.645	0.706	0.650	0.434	-0.790	-0.05	0.416
k^q	0.996	1.001	0.993	1.003	0.999	0.996	0.998
$(R^2 - R_0^2)/R^2$	0.275	0.206	0.228	0.503	2.406	1.066	0.486

q^2 , PRESS, N, Noc, r^2 , S, F, estérico y electrostático tienen el mismo significado que en la Tabla 13.

^{j,k}Contribución del parámetro donador y aceptor de hidrógeno respectivamente.

^lContribución del parámetro hidrófobo.

^mCoefficiente de correlación para el grupo de prueba.

^{n,p}Coefficiente de correlación para la ordenada al origen, para las gráficas actividad experimental & actividad calculada y actividad calculada & actividad experimental, usando el grupo de prueba.

^{o,q}Pendientes para la ordenada al origen de las curvas actividad calculada & actividad experimental y actividad experimental & actividad calculada.

El modelo 6 se construyó empleado las cargas derivadas del potencial electrostático calculadas usando el método semiempírico PM3. Este modelo presentó alta correlación y valor de q^2 ($r^2 = 0.936$ y $q^2 = 0.634$, Tabla 15). Los resultados indican que la fracción electrostática muestra un valor importante (0.475); por esta razón se decidió explorar otro tipo de cargas atómicas. Los modelos 7 y 8 se construyeron usando las cargas de Mulliken y Gasteiger-Hückel respectivamente. Estos modelos muestran alta correlación ($r^2 = 0.927$ y 0.928 , respectivamente) y buenos valores de q^2 (0.602 y 0.627, respectivamente); no obstante, el modelo 6 presentó los mejores valores. Para generar los modelos de CoMSIA 9–12 se emplearon los mismos conformeros y el mismo tipo de cargas usadas en el modelo 6. El modelo 9 se construyó usando los parámetros estérico y electrostático, el modelo 10 incluye los

parámetros donador y aceptor de hidrógeno y el modelo 11 el hidrófobo. Los modelos de CoMSIA 9–11 muestran valores aceptables de r^2 y q^2 ; particularmente, el modelo 9 mostró los valores más altos ($r^2 = 0.915$ y $q^2 = 0.617$). El modelo 12 se construyó usando todos los parámetros mencionados y el valor de q^2 incrementó con respecto a los modelos de CoMSIA previos. Aunque las diferencias entre los modelos 9 y 12 son pequeñas, los mapas de contornos son más informativos en el modelo que contiene los cinco parámetros.

7.2.3. Modelos construidos con conformaciones selectas empleando similitud 3D

Los modelos generados a partir de las conformaciones de mínima energía están restringidos a que todos los compuestos muestren un arreglo tridimensional común; sin embargo, se puede definir un espacio restringido basado en que la unión de moléculas a un receptor se caracteriza por una alta complementariedad e interacciones específicas [27]. En este trabajo se utilizó la metodología de superposición rápida de estructuras químicas implementada en el programa ROCS para calcular la similitud entre moléculas basada en su arreglo tridimensional [68]. Esta metodología se utilizó por primera vez en este trabajo para seleccionar confórmeros en estudios QSAR-3D. Se seleccionaron tres moléculas como referencia en su conformación de mínima energía. La primera de ellas, **45**, se seleccionó por tener el sustituyente más largo en la posición 2 del núcleo de bencimidazol y adicionalmente por ser una de las moléculas más activas. Las moléculas **9** y **10** se seleccionaron considerando su cadena larga y fija por efecto de tautomería en las posiciones 5 y 6, respectivamente. Para el resto de los compuestos con enlaces rotables, los confórmeros se seleccionaron mediante superposición y tomando en cuenta el mayor valor de similitud (combo score) con la molécula de referencia correspondiente. Se observó que todas las moléculas sobrepuestas usando ROCS muestran un alineamiento casi perfecto en el núcleo de bencimidazol, por esto se decidió considerar el alineamiento basado en el bencimidazol para todos los compuestos en los modelos QSAR (Figura 32).

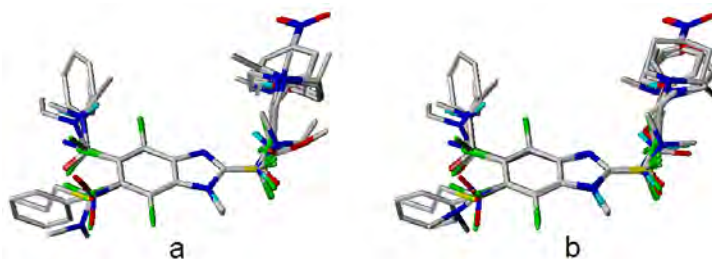


Figura 32. Moléculas incluidas en el grupo de trabajo usando (a) las conformaciones de mínima energía y (b) Las conformaciones seleccionadas por superposición con **9**, **10** y **45**.

El modelo de CoMFA 13 se construyó con confórmeros selectos por similitud en 3D, las cargas derivadas del potencial electrostático y los tautómeros definidos por el modelo 2. Se observó alta correlación ($r^2 = 0.934$) y un alto valor del coeficiente LOO ($q^2 = 0.601$, Tabla 16).

Tabla 16. Resumen de resultados para los modelos de CoMFA 13–17^a

	Modelo CoMFA		Modelos CoMSIA		
	13	14	15	16	17
q^2	0.601	0.565	0.482	0.506	0.585
PRESS	0.489	0.510	0.551	0.538	0.484
N	55	55	55	55	55
Noc	5	5	4	4	2
r^2	0.934	0.914	0.669	0.748	0.805
S	0.198	0.227	0.441	0.385	0.332
F	139.397	104.593	25.296	37.061	107.175
Estérico	0.499	0.245			0.095
Electrostático	0.501	0.755			0.307
Donador			0.603		0.247
Aceptor			0.397		0.121
Hidrofóbico				1	0.230
R^2	0.825	0.729	0.689	0.678	0.760
R_0^2	0.795	0.692	0.589	0.644	0.726
k	0.995	0.993	0.996	1.009	0.995
$(R^2 - R_0^2)/R^2$	0.036	0.051	0.145	0.050	0.045
$R_0'^2$	0.625	0.301	-0.54	0.135	0.426
k'	1.002	1.002	0.998	0.986	1.001
$(R^2 - R_0'^2)/R^2$	0.242	0.587	1.784	0.801	0.439

^aTodos los valores tienen el mismo significado que en las tablas 13 y 15.

Los resultados para el modelo 13 son comparables con los modelos de CoMFA basados en conformaciones de mínima energía. Adicionalmente, se construyeron los modelos de CoMSIA 14–17, empleando las mismas estructuras y cargas atómicas que el modelo anterior. Los cuatro modelos generados incluyen, al igual que los modelos 9–12, los parámetros estérico, electrostático, donador, aceptor, hidrófobo y la

combinación de ellos, respectivamente. Los resultados mostraron que el modelo 14, que incluye los parámetros estéricos y electrostáticos, y el modelo 17, que incluye la combinación de los cinco parámetros, tienen los mejores valores de r^2 y q^2 .

Los modelos basados en selección de confórmeros por similitud 3D muestran resultados muy similares a los modelos obtenidos con confórmeros de mínima energía; por lo que pueden ser utilizados cuando los modelos 6–12 fallen en la predicción de moléculas que contienen varios enlaces rotables.

7.2.4. Validación

La validación es una parte esencial en los estudios de relaciones cuantitativas estructura–actividad. Normalmente se utiliza la validación cruzada como criterio para definir la capacidad predictiva de los modelos. Sin embargo, se ha demostrado que los modelos con valores altos de LOO q^2 (> 0.5) no siempre son predictivos [49–50]. Con el fin de tener una mejor validación, es necesario contar con un grupo de prueba que contenga un mínimo de 5 compuestos cuyo espacio químico y rango de actividades sean comparables con el grupo de trabajo (moléculas usadas en la construcción del modelo) [30, 49–50]. En este trabajo se utilizaron 13 moléculas como grupo de prueba y sus actividades se calcularon usando los modelos 6–17. Adicionalmente, se construyeron las curvas de regresión de actividad experimental contra actividad predicha y actividad predicha contra actividad experimental para aplicar las condiciones definidas en la sección 6.2.9. La primera y segunda condición se refieren a que el coeficiente de validación LOO q^2 y el coeficiente de correlación para el grupo de prueba R^2 deben ser al menos 0.5 y 0.6, respectivamente. Estas condiciones son necesarias pero no determinantes para definir un modelo como predictivo. La tercera condición ($[(R^2 - R_0^2)/R^2] < 0.1$ ó $[(R^2 - R'_0{}^2)/R^2] < 0.1$) se refiere a que al menos uno de los coeficientes de correlación de la línea de regresión que pasa por el origen ($R'_0{}^2$ ó R_0^2) debe ser cercano a la correlación de la línea de tendencia para las predicciones (R^2); mientras que la cuarta condición ($0.85 \leq k \leq 1.15$ ó $0.85 \leq k' \leq 1.15$) se refiere a que al menos una de las pendientes de las líneas de regresión que pasan por el origen debe ser cercana a 1. La aplicación de estas condiciones sugiere que los modelos se acercan a ser modelos “ideales” (predictivos) [49]. Los valores obtenidos de los

gráficos de regresión y la aplicación de las condiciones I–V se muestran en las Tablas 15 y 16. En la Figura 33 se muestra un ejemplo de los gráficos de regresión para el modelo 6.

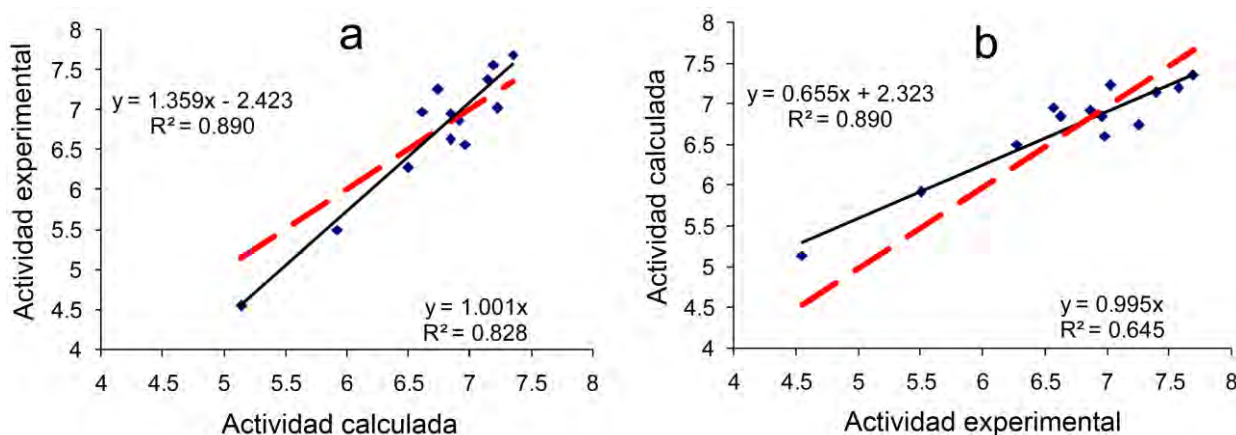


Figura 33. Validación del modelo de CoMFA 6. (a) Actividad experimental contra actividad calculada, (b) actividad calculada contra actividad experimental. La línea de tendencia se muestra con un trazo continuo mientras que la línea de regresión que pasa por el origen se muestra con un trazo discontinuo.

Los modelos de CoMFA 6–8 satisfacen las condiciones de validación I–IV; particularmente el modelo 6 mostró los mejores valores de q^2 y R^2 . En este caso, el modelo 6 coincide en que al tener el mejor valor de correlación para la validación interna se obtiene el mejor valor de correlación para la validación externa. Por otro lado, los modelos de CoMSIA 9–12 muestran valores aceptables de q^2 ; en especial el modelo 12 mostró el valor más alto de q^2 y el segundo valor más alto de R^2 . Sin embargo, los modelos 9 y 11 no cumplen la tercera condición y el modelo 10 no cumple la segunda; por esta razón, sólo el modelo 12 se considera predictivo. Los modelos basados en confórmeros elegidos por similitud 3D satisfacen las condiciones I–IV, con excepción del modelo 15. Aunque los modelos 13–17 tienden a mostrar valores más bajos de q^2 y R^2 comparados con los modelos 6–12, cumplen mejor con la tercera condición.

Los valores calculados para el grupo de prueba usando los mejores modelos (6, 12, 13 y 17) se muestran en la Tabla 17.

Tabla 17. Predicciones para el grupo de prueba usando los modelos 6, 12, 13 y 17

	Modelo 6			Modelo 12		Modelo 13		Modelo 17	
	Exp. ^a	Calc. ^b	Res. ^c	Calc.	Res.	Calc.	Res.	Calc.	Res.
3a	5.50	5.92	-0.42	6.18	-0.68	5.88	-0.38	6.09	-0.59
11a	4.55	5.14	-0.59	5.4	-0.85	5.17	-0.62	5.33	-0.78
22b	7.03	7.23	-0.2	7.32	-0.29	7.24	-0.21	7.16	-0.13
27a	6.57	6.96	-0.39	6.71	-0.14	6.89	-0.32	6.56	0.01
36a	7.39	7.14	0.25	7.36	0.03	7.41	-0.02	7.51	-0.12
41a	7.69	7.35	0.34	7.52	0.17	7.28	0.41	7.42	0.27
46	6.96	6.84	0.12	6.81	0.15	7.04	-0.08	6.91	0.05
48	6.63	6.84	-0.21	6.97	-0.34	6.8	-0.17	7.25	-0.62
53	7.57	7.19	0.38	7.09	0.48	7.26	0.31	7.13	0.44
54	6.87	6.91	-0.04	6.81	0.06	7.18	-0.31	6.92	-0.05
59	6.98	6.61	0.37	6.74	0.24	6.3	0.68	6.77	0.21
65	6.27	6.50	-0.23	6.25	0.02	6.62	-0.35	6.46	-0.19
69a	7.26	6.74	0.52	6.53	0.73	6.8	0.46	6.38	0.88

^aActividad experimental.

^bActividad calculada.

^cResidual.

7.2.5. Mapas de contornos de CoMFA y CoMSIA

Los mapas de contornos para los mejores modelos de CoMFA y CoMSIA (6 y 12, respectivamente) se muestran en la Figura 34. La discusión sobre los mapas se muestra en las siguientes seis secciones.

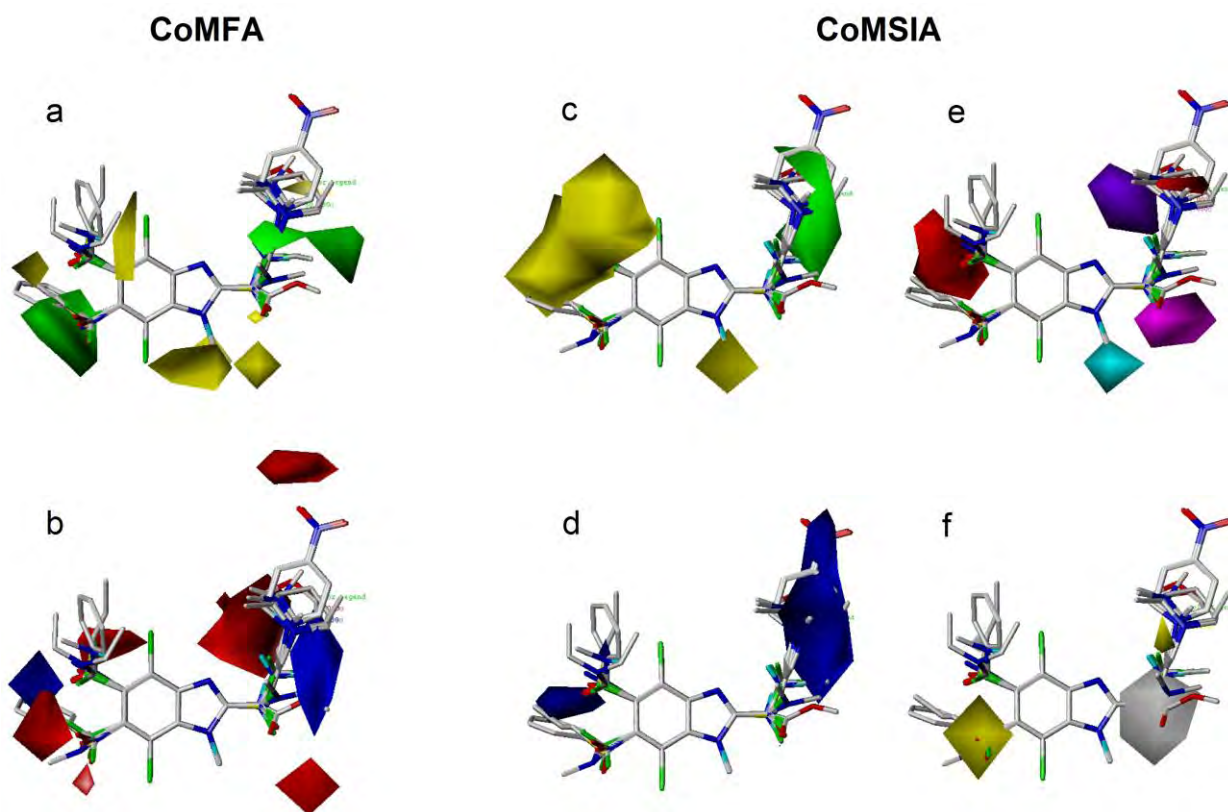


Figura 34. Mapas de contornos para los mejores modelos. (CoMFA) Mapas de contornos estéricos y electrostáticos para el modelo 6. (a) Las regiones donde se incrementa la actividad con sustituyentes voluminosos están en verde y donde se disminuye en amarillo. (b) Las regiones donde la actividad se favorece por cargas atómicas parciales negativas están en rojo y donde se desfavorece están en azul. (CoMSIA) Mapas de contornos estéricos, electrostáticos, donador-aceptor e hidrófobos para el modelo 12. (c), (d) Los contornos tienen el mismo significado que en el modelo de CoMFA. (e) Los sitios donde la actividad se favorece con donadores y aceptores de hidrógeno están en color cian y magenta respectivamente. Los sitios donde la actividad se desfavorece por donadores y aceptores están en morado y rojo respectivamente. (f) Los sitios donde sustituyentes hidrófobos favorecen la actividad se muestran en amarillo y donde la actividad se desfavorece en gris.

7.2.5.1. Mapas de contorno estéricos para el mejor modelo de CoMFA

Los mapas de contorno estéricos (Figura 34a) muestran en color verde las regiones donde la actividad se favorece cuando aumenta el volumen del sustituyente y en amarillo las regiones donde la actividad disminuye cuando aumenta el volumen del sustituyente.

En la posición 1 del núcleo de bencimidazol se observa un contorno amarillo señalando que la actividad se desfavorece en los compuestos sustituidos con metilo, en general, esta tendencia se observa para todos los derivados de bencimidazol con sustituyentes pequeños (1, 2, 4–7, 14–21, 23–26 y 29–30) donde los 1-

metilbencimidazoles **6** y **7** son los compuestos menos activos ($pCl_{50} = 5.39$ y 5.27 , respectivamente). Por otra parte, los compuestos más activos ($pCl_{50} > 7$) tienen hidrógeno en la posición 1 del bencimidazol con excepción de los compuestos **49**, **56**, **57** y **61**, en los cuales el sustituyente en posición 2 parece jugar un papel muy importante en la actividad. Esta tendencia es diferente para los compuestos **8–10**, **12** y **13**, en los cuales los sustituyentes largos en las posiciones 5 y 6 determinan la diferencia en la actividad. El contorno verde cercano a la posición 2 del bencimidazol puede asociarse con la alta actividad de los derivados de 2-alquiltiobencimidazol **32–35**, **37–40**, **42–45**, y los derivados con 2-carboxi **49**, **57**, y **61** (pCl_{50} 6.16–8.48). Este último grupo contiene la mayoría de los compuestos más activos de la base de datos. Los contornos color amarillo y verde, cercanos a la posición 5 y 6 respectivamente, sugieren que los compuestos sustituidos en posición 6 serán más activos que aquellos sustituidos en la posición 5. Este comportamiento es observado en los derivados de 1-metilbencimidazol **12**, **13**, **47**, **51**, **63** y **64**. Sin embargo, esta tendencia no es muy clara en los compuestos **6** y **7** porque tienen casi la misma actividad. Por otra parte, para los compuestos **9** y **10** se observa la tendencia opuesta. La diferencia en la actividad para los últimos dos compuestos, puede explicar el contorno en el borde de la posición 6, sugiriendo que los sustituyentes largos disminuyen la actividad.

7.2.5.2. Mapas de contorno electrostáticos para el mejor modelo de CoMFA

Los mapas de contorno electrostáticos se muestran en rojo cuando la actividad se favorece con cargas atómicas parciales negativas y en azul cuando la actividad se desfavorece con cargas atómicas parciales negativas o se favorece con positivas (Figura 34b). En general, los contornos rojos se muestran cercanos a heteroátomos como nitrógenos y oxígenos, en los cuales las cargas atómicas parciales calculadas tienden a ser más negativas. Se observan tres contornos rojos sobre los sustituyentes en la posición 2. El primero de ellos se encuentra cercano a los oxígenos sp^2 en los compuestos **47**, **49–52**, **55–58**, y **60–61**, que muestran actividad moderada para los derivados de tipo amida (pCl_{50} 6.37–7.12) y alta actividad para los derivados con un grupo 2-carboxi (pCl_{50} 7.07–7.72). Un segundo contorno rojo se encuentra sobre posición 2 cercano a los nitrógenos de tipo amino en los compuestos **34**, **37–40**, y **42–**

43 los cuales muestran actividad alta ($pCl_{50} > 7$). El último contorno rojo se encuentra sobre el extremo de la posición 2 y se asocia con los oxígenos del grupo nitro en **44** y **45** que son dos de los compuestos más activos ($pCl_{50} = 7.96$ y 8.48 , respectivamente). Un contorno azul sobre la posición 2 señala a la porción alifática de los compuestos **32–35**, **37–40**, y **42–45**, los cuales presentan actividad que va desde moderada a alta ($pCl_{50} = 6.16–8.48$). El mismo contorno se encuentra cercano al etilo de los derivados de ésteres **49**, **57**, y **61** ($pCl_{50} > 7$) que muestran una actividad alta. Sobre las posiciones 5 y 6 se encuentran tres contornos rojos que sugieren que los sustituyentes como los halógenos incrementan la actividad comparados con sus análogos sustituidos con hidrógeno. Además, los sustituyentes largos son asociados con un contorno azul cercano a las posiciones 5 y 6, lo que sugiere un efecto favorable asociado a la porción alifática de los compuestos con sustituyente propiltio **8–10** ($pCl_{50} = 5.59–6.70$) y a las alquilamidas **68** y **70** ($pCl_{50} = 6.05–7.54$) en comparación con el benzoilo en los compuestos menos activos **12–13** ($pCl_{50} = 4.53$ y 4.97 , respectivamente).

7.2.5.3. Mapas de contorno estéricos para el mejor modelo de CoMSIA

El código de colores en los mapas de contornos estéricos de CoMSIA (Figura 34c), tienen el mismo significado que en los mapas de CoMFA. Los contornos de CoMSIA son similares a los de CoMFA; sin embargo, son más definidos. Sobre la posición 1 se observa un contorno amarillo, que al igual que en CoMFA, sugiere el efecto favorable del hidrógeno. Sobre la posición 2 se observa un contorno verde que sugiere un efecto favorable en los derivados con sustituyentes largos en esta posición. De manera similar que en CoMFA, se observa un contorno amarillo sobre la posición 5 y en el extremo de la posición 6, indicando el efecto desfavorable de los sustituyentes largos en las posiciones 5 y 6.

7.2.5.4. Mapas de contorno electrostáticos para el mejor modelo de CoMSIA

El color de los contornos electrostáticos en los modelos CoMSIA (Figura 34d) tienen el mismo significado que en los modelos de CoMFA. De manera similar a los contornos electrostáticos de CoMFA, se observa un contorno azul sobre la posición 2

del bencimidazol; sin embargo, el contorno de CoMSIA es más grande y se extiende sobre la porción alifática de los compuestos **32–35**, **37–40** y **42–43**, y sobre el anillo aromático desactivado por efecto del grupo nitro en los compuestos **44** y **45**. Adicionalmente, al igual que en CoMFA, un segundo contorno azul se observa cercano a las posiciones 5 y 6. Este contorno está cercano al anillo aromático del compuesto **13**, el cual tiene actividad pobre ($pCl_{50} = 4.97$), y a los carbonilos de los compuestos **12**, **66**, **67** y **70** ($pCl_{50} = 4.53–6.79$, mediana $pCl_{50} = 6.36$) que tienen actividad que va de pobre a moderada.

7.2.5.5. Mapas de contorno de donadores y aceptores de hidrógeno para el mejor modelo de CoMSIA

Las regiones favorecidas por donadores y aceptores de hidrógeno se muestran en color ciánico y magenta, respectivamente; mientras que las áreas desfavorecidas se muestran en morado y rojo, respectivamente (Figura 34e). Sobre la posición 1 del bencimidazol se observa un contorno ciánico que sugiere la contribución favorable del hidrógeno en los derivados de bencimidazol; este contorno coincide con los contornos estéricos color amarillo de CoMFA y CoMSIA generados por los derivados metilados. Sobre la posición 2, se observa un contorno morado el que sugiere el efecto desfavorable de donadores de hidrógeno; la aparición de este contorno se asocia con el decremento en la actividad en los compuestos hidroxilados **32** y **33** ($pCl_{50} < 7$) y con los N-H en los compuestos de tipo amida que son menos activos que sus análogos de tipo éster. Un contorno magenta se observa en la posición 2 señalando directamente a los oxígenos sp^2 en los compuestos de tipo amida y éster que pueden ser importantes como aceptores de hidrógeno. Adicionalmente, se observa un contorno rojo cercano a la posición 2 del bencimidazol que puede asociarse con el incremento en la actividad cuando la cadena que conecta al bencimidazol con el aceptor de hidrógeno cambia de dos a tres metilenos en los compuestos **34–35** y **38–39** (pCl_{50} 6.51–7.14 a 7.07–7.99, respectivamente). Un segundo contorno rojo se observa cercano a los oxígenos sp^2 de los derivados con sustituyentes amida y benzoilo en la posición 5.

7.2.5.6. Mapas de contorno hidrófobos para el mejor modelo de CoMSIA

Los mapas de contornos hidrófobos se muestran en amarillo cuando la actividad aumenta con sustituyentes hidrófobos y en gris cuando la actividad disminuye (Figura 34f). Se observa un contorno gris en la posición 2 del bencimidazol generado por los sustituyentes trifluorometilo y pentafluoroetilo que presentan actividad desde pobre a moderada (pCl_{50} 4.53–6.98, mediana pCl_{50} = 6.24). Adicionalmente, sobre la posición 2 se encuentra un contorno amarillo que señala a la porción alifática de los compuestos **38–39**. Un segundo contorno amarillo se encuentra cercano a la posición 6, este contorno puede explicarse por la presencia de halógenos que en la mayoría de los casos producen compuestos un poco más activos particularmente en esta posición. Este último contorno es similar y complementario en información al observado en CoMFA.

7.3. Actividad antiprotozoaria de los inhibidores de la bomba de protones

7.3.1. Inhibidores de la bomba de protones

Los derivados de bencimidazol son compuestos con una amplia variedad de actividades biológicas reportadas y en algunos casos éstos se usan en la terapéutica [89]. Por ejemplo, los derivados de bencimidazol conocidos como inhibidores de la bomba de protones (PPIs) son ampliamente usados para el tratamiento de úlcera péptica, reflujo gastroesofágico y erradicación de *Helicobacter pylori* (en combinación con antibióticos) [90]. Basados en las similitudes estructurales de los PPIs con algunos derivados de bencimidazol con actividad antiprotozoaria y en los estudios de relaciones estructura–actividad presentados en las secciones anteriores, se evaluó su actividad contra algunos protozoarios.

7.3.2. Similitud entre PPIs y algunos bencimidazoles antiprotozoarios

El núcleo de bencimidazol es la estructura base de muchos derivados con potente actividad contra *Trichomonas vaginalis*, *Giardia intestinalis* y *Entamoeba histolytica* y para los PPIs (Figura 35). Adicionalmente, los sustituyentes juegan un papel importante en la actividad biológica.

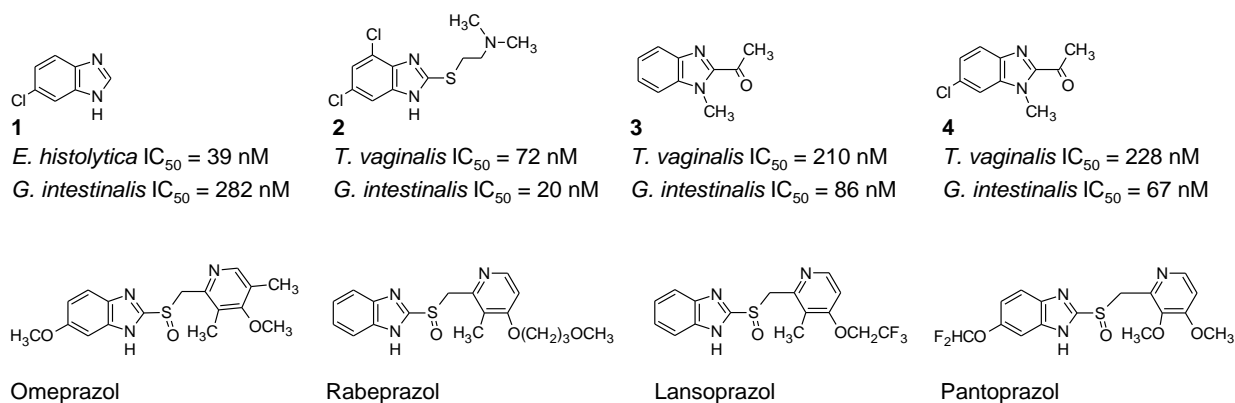


Figura 35. Algunos ejemplos de derivados de bencimidazol con actividad antiprotozoaria y los PPIs comerciales.

Para visualizar la similitud en 3D entre los compuestos **1–4** y los PPIs, se obtuvieron confórmeros de mínima energía usando modelado molecular. Posteriormente, las moléculas se alinearon tomando en cuenta el núcleo de

bencimidazol y aceptores de hidrógeno comunes sobre la posición 2 del bencimidazol. Los resultados mostraron que los PPIs pueden adquirir un arreglo tridimensional similar a los derivados **1–4** ubicando los posibles aceptores de hidrógeno en posiciones similares como se muestra en la Figura 36. Es importante notar que las distancias entre el bencimidazol (centroide localizado en el anillo de benceno del bencimidazol) y los aceptores de hidrógeno comunes sobre la posición 2 son muy similares entre las moléculas de referencia y los PPIs.

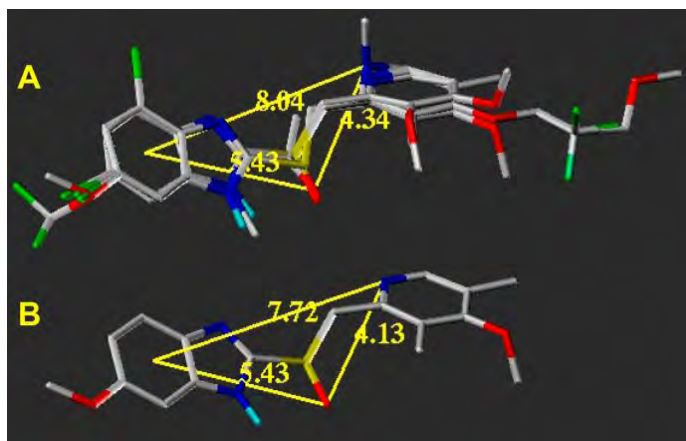


Figura 36. (A) Sobreposición de los compuestos de referencia **1–4** y los PPIs. Las distancias tipo triángulo fueron medidas con las moléculas **2** y **3** considerando algunos de sus posibles elementos farmacofóricos. (B) Distancias tipo triángulo para el omeprazol.

Adicionalmente se calcularon los mapas de potencial electrostático usando las mismas estructuras. Los resultados muestran una alta similitud electrónica entre las moléculas de referencia y los PPIs (Figura 37). Los isocontornos cercanos a la posición 3 del bencimidazol son comunes en todos los derivados, además, los isocontornos cercanos a los oxígenos sp^2 y nitrógenos sobre la posición 2 son comunes para algunas moléculas de referencia y los PPIs. Por otra parte, los compuestos **1**, **2** y **4** muestran isocontornos similares al omeprazol y pantoprazol sobre la posición 5(6).

Estas observaciones, y los resultados previos para derivados de bencimidazol sobre el panorama de actividad y QSAR-3D, sugerían que los PPIs podrían tener actividad antiprotozoaria por un mecanismo similar a los compuestos **1–4**, el cual se desconoce actualmente.

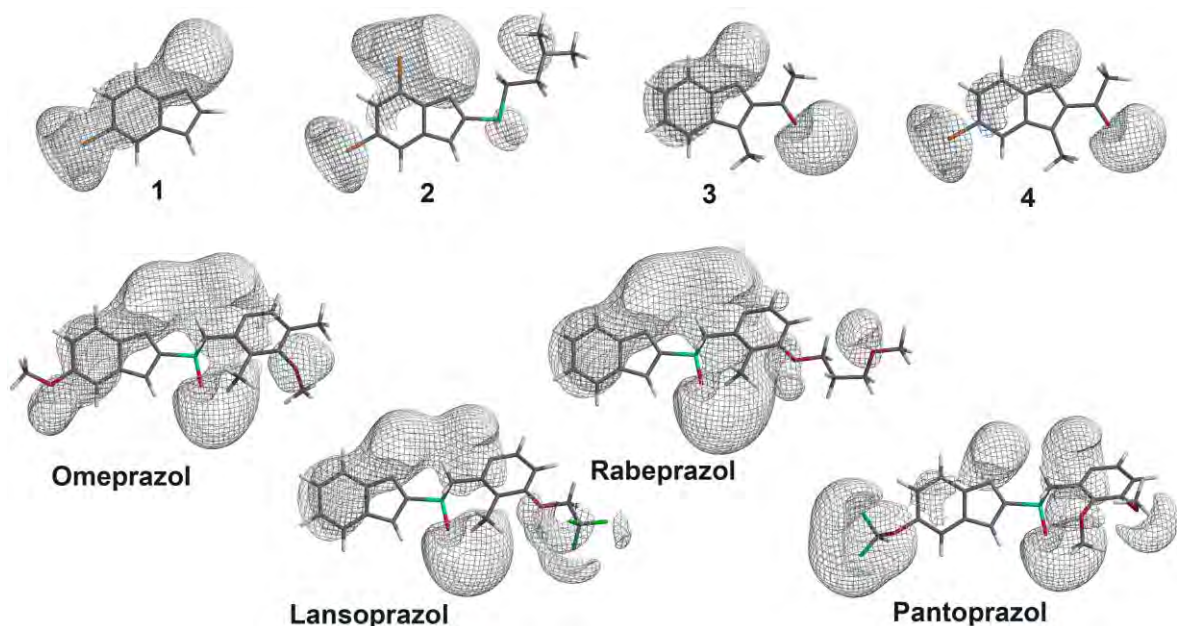


Figura 37. Potencial electrostático molecular para los derivados 1–4 y los inhibidores de la bomba de protones calculados usando el método de Hartree-Fock (HF/6-31G**/HF/6-31G*) a un isovalor de -10 kcal/mol.

7.3.3. Evaluación antiprotozoaria de los PPIs

Los resultados de la actividad antiprotozoaria (in vitro) para los PPIs se muestran en la Tabla 18. Todos los compuestos mostraron potente actividad antiprotozoaria con valores en el rango nanomolar. Los PPIs fueron de 1.9 a 3.1 veces más activos que el metronidazol contra *T. vaginalis* y 12.8 a 78.1 veces más activos contra *G. intestinalis*. Adicionalmente, el rabeprazol y el pantoprazol fueron 14.8 y 134.6 veces más activos contra *E. histolytica* respectivamente. Es importante notar que el pantoprazol mostró la mejor actividad contra los tres protozoarios evaluados.

Tabla 18. Actividad in vitro de los PPIs comerciales contra *T. vaginalis*, *G. intestinalis* y *E. histolytica*

	<i>T. vaginalis</i> Cl_{50} μ M	<i>G. intestinalis</i> Cl_{50} μ M	<i>E. histolytica</i> Cl_{50} μ M
Omeprazol	0.1216 ± 0.0058	0.0955 ± 0.0029	0.4922 ± 0.0289
Lansoprazol	0.1218 ± 0.0027	0.0731 ± 0.0027	0.3466 ± 0.0190
Rabeprazol	0.1057 ± 0.0056	0.0181 ± 0.0014	0.0237 ± 0.0014
Pantoprazol	0.0756 ± 0.0026	0.0157 ± 0.0000	0.0026 ± 0.0000
Albendazol	1.5905 ± 0.0113	0.0370 ± 0.0030	56.5334 ± 18.8445
Metronidazol	0.2360 ± 0.0160	1.2260 ± 0.1250	0.3798 ± 0.1461

Aunque el omeprazol y el lansoprazol tienen diferentes sustituyentes en el núcleo de bencimidazol y el anillo de piridina, no hay diferencia importante en su actividad contra *T. vaginalis* y las diferencias contra *G. intestinalis* y *E. histolytica* son muy pequeñas. Las diferencias en actividad son más notorias cuando se comparan el lansoprazol y el rabeprazol, sugiriendo que el sustituyente en la posición 2, del núcleo de bencimidazol, juega un papel muy importante en la actividad. Además, el incremento en la actividad en el pantoprazol comparado con el omeprazol contra los tres protozoarios sugiere que existe una contribución electrostática importante causada posiblemente por el grupo difluoroetoxi en posición 5(6) del núcleo de bencimidazol.

Considerando la nueva actividad antiprotozoaria para los PPIs, estos compuestos son candidatos para reposicionamiento de fármacos; esto es, un nuevo uso. Esto bajo el hecho de que tienen una farmacocinética y seguridad conocida para su uso en humanos [91].

El reposicionamiento de fármacos es una estrategia popular en compañías farmacéuticas debido a que lleva un tiempo relativamente corto entre los ensayos de laboratorio y la clínica. Por lo tanto, el reposicionamiento es una excelente estrategia para acceder a nuevas terapias; especialmente para las enfermedades poco atendidas como las parasitarias [91–94].

7.4. Síntesis de nuevos derivados de bencimidazol con actividad antiprotozoaria

7.4.1. Diseño de las moléculas

Uno de los aspectos más importantes en el diseño de fármacos es la optimización de estructuras líder. Para esto, el diseño racional de nuevos derivados y el estudio de relaciones estructura–actividad juegan un papel muy importante. A partir de los estudios SAR y QSAR realizados anteriormente, contra *T. vaginalis* y *G. intestinalis*, se tomaron en cuenta varias conclusiones importantes para el diseño de los nuevos derivados del bencimidazol sintetizados en este trabajo. Adicionalmente, para el caso de *E. histolytica*, estudios de CoMFA previos sugieren algunas características importantes que los bencimidazoles deben tener para mejorar su actividad antiamebiana [48]; sin embargo, dado a que las moléculas en dicho estudio poseen pocas variaciones estructurales, hace falta enriquecer la base de datos de bencimidazoles para saber más acerca de las relaciones estructura–actividad antiamebiana [48].

Algunas de las conclusiones más importantes generadas por los estudios anteriores se resumen en la Figura 38.

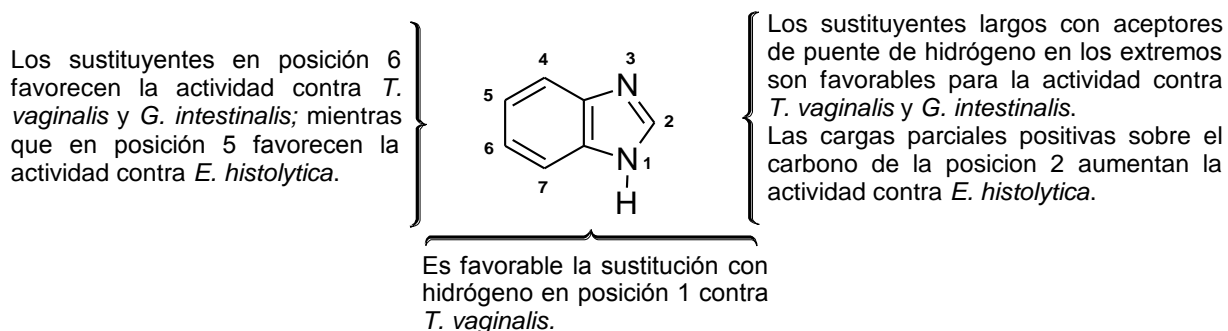


Figura 38. Conclusiones más importantes sobre relaciones estructura-actividad para bencimidazoles evaluados contra *T. vaginalis*, *G. intestinalis* y *E. histolytica*.

La primera consideración se tomó a partir de la conclusión sobre el efecto favorable con grupos grandes en posición 2 como lo sugieren los modelos de CoMFA para *T. vaginalis* (Figura 34) y *G. intestinalis* (Figura 13). Es importante notar que algunos de los compuestos más activos contra ambos protozoarios tienen sustituyentes grandes en posición 2 con aceptores de hidrógeno en el extremo del

sustituyente o un anillo aromático; teniendo en cuenta esto, se diseñó una serie de moléculas y se sometieron a cribado virtual. El cálculo de la actividad se realizó usando los mejores modelos de CoMFA y CoMSIA construidos para *T. vaginalis*. Aunque en ese momento no se contaba con el modelo construido para *G. intestinalis*, sólo se tomó en cuenta que los mapas de contornos de los modelos de CoMFA son muy similares contra ambos protozoarios. De los resultados obtenidos, el compuesto **S1J01** (Figura 39) fue interesante desde el punto de vista de su actividad calculada y su facilidad de preparación. Este compuesto se diseñó tomando en cuenta dos compuestos muy activos y combina un anillo aromático sobre la posición 2 con un aceptor de hidrógeno (Figura 39).

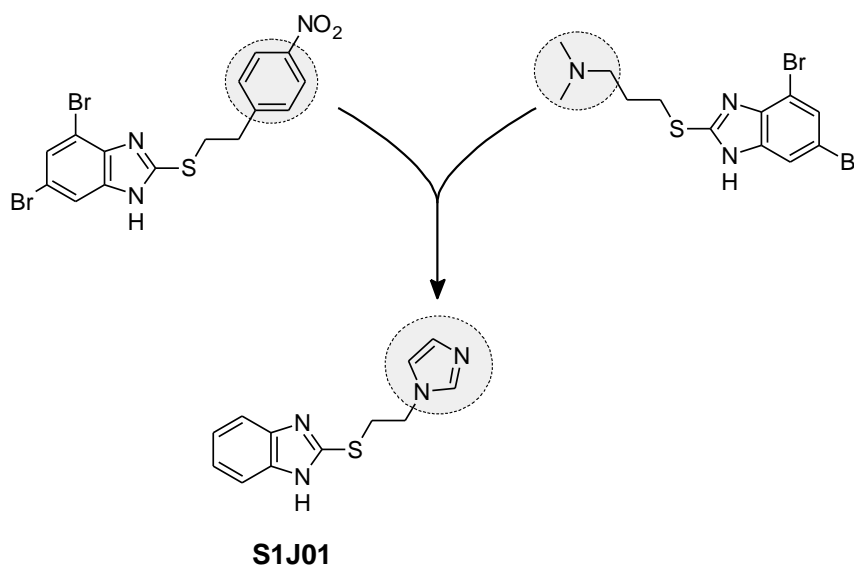


Figura 39. Consideraciones para el diseño del compuesto **S1J01**.

Con la estructura base del compuesto **S1J01** se diseñaron 18 derivados por modificaciones sobre las posiciones 1, 5 y 6. Los nuevos compuestos contienen hidrógeno o metilo en posición 1, esto se consideró teniendo en cuenta que la actividad se favorece con hidrógeno contra *T. vaginalis*; sin embargo, el grupo metilo podría mejorar su lipofilia sin una baja importante en la actividad. Finalmente, en las posiciones 5 y 6 se consideraron sustituyentes cloro, etoxi y éster metílico, tomando en cuenta que las sustituciones en posición 6 favorecen la actividad contra *T. vaginalis* y *G. intestinalis*; y en posición 5 contra *E. histolytica*. Debido a que aún es poca la información sobre el efecto de los sustituyentes en estas posiciones es importante

generar nuevos derivados. Los resultados de actividad calculada, empleando los mejores modelos QSAR, para los 19 compuestos diseñados se muestran en la Tabla 19.

Tabla 19. Actividad calculada para las moléculas diseñadas

Clave	R ₁	R ₂	R ₃	Modelo 6	Modelo 12	Modelo 13	Modelo 17
S1J01	H	H	H	6.82	7.04	6.59	7.07
S1J02	H	Cl	H	(a) 6.95, (b) 7.25	(a) 7.00, (b) 7.29	(a) 6.74, (b) 7.05	(a) 7.05, (b) 7.31
S1J03	H	Cl	Cl	7.31	7.25	7.21	7.28
S1J04	CH ₃	H	H	6.49	6.53	6.29	6.74
S1J05	CH ₃	H	Cl	6.89	6.79	6.72	6.98
S1J06	CH ₃	Cl	H	6.55	6.50	6.41	6.72
S1J07	CH ₃	Cl	Cl	6.93	6.76	6.84	6.95
S2J08	H	COOCH ₃	H	(a) 6.69, (b) 6.93	(a) 6.49, (b) 6.80	(a) 6.62, (b) 6.88	(a) 6.52, (b) 6.93
S2J09	H	COOCH ₃	Cl	(a) 7.28, (b) 7.30	(a) 6.82, (b) 6.90	(a) 7.16, (b) 7.42	(a) 6.85, (b) 6.98
S2J10	CH ₃	H	COOCH ₃	6.56	6.29	6.58	6.59
S2J11	CH ₃	Cl	COOCH ₃	7.27	6.57	7.24	6.69
S2J12	CH ₃	COOCH ₃	H	6.40	6.00	6.30	6.21
S2J13	CH ₃	COOCH ₃	Cl	6.93	6.33	6.81	6.53
S3J14	H	OCH ₂ CH ₃	H	(a) 7.02, (b) 7.16	(a) 6.98, (b) 7.05	(a) 6.90, (b) 7.05	(a) 7.02, (b) 7.08
S3J15	H	OCH ₂ CH ₃	Cl	(a) 7.34, (b) 7.25	(a) 7.23, (b) 7.13	(a) 7.34, (b) 7.21	(a) 7.25, (b) 7.14
S3J16	CH ₃	H	OCH ₂ CH ₃	6.74	6.54	6.74	6.75
S3J17	CH ₃	Cl	OCH ₂ CH ₃	6.85	6.64	6.87	6.81
S3J18	CH ₃	OCH ₂ CH ₃	H	6.59	6.43	6.60	6.67
S3J19	CH ₃	OCH ₂ CH ₃	Cl	6.94	6.69	7.01	6.91

*(a) y (b) representan la actividad para las moléculas con patrones de sustitución en relación 1–5 y 1–6 respectivamente en base al núcleo de benzimidazol. **La actividad está dada en $-\log Cl_{50}$ donde la Cl_{50} está dada en unidades molares.

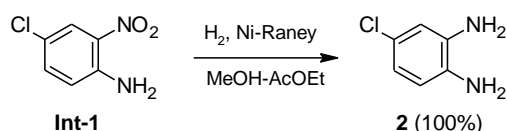
7.4.2. Síntesis de los 1,2-Diaminobencenos (2–17)

Los 1,2-diaminobencenos se prepararon a partir de derivados de benceno comerciales (con excepción del 1,2-diaminobenceno **1** que estaba disponible comercialmente). Los resultados para cada una de las rutas se discuten en las siguientes secciones.

7.4.2.1. 4-Cloro-1,2-diaminobenceno (2)

La síntesis del compuesto **2** se llevó a cabo en un solo paso a partir de la 4-cloro-2-nitroanilina comercial mediante una reacción de hidrogenación catalítica usando Ni-Raney como catalizador (Esquema 6). Es importante resaltar la elección del

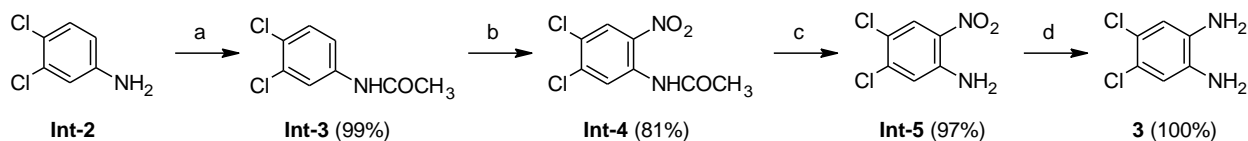
catalizador; en general, se usó Pd/C para materias primas no cloradas y Ni-Raney para las materias primas cloradas debido a que el Pd/C puede generar productos secundarios de reacción. Es importante mencionar que la mayoría de los 1,2-diaminobencenos obtenidos en este trabajo tienden a oxidarse al contacto con el aire; por esta razón, se utilizaron inmediatamente en la siguiente reacción.



Esquema 6. Síntesis del compuesto 2.

7.4.2.2. 4,5-Dicloro-1,2-diaminobenceno (3)

Para la síntesis de **3** se partió de la 3,4-dicloroanilina comercial (**Int-2**) y se siguió una secuencia de cuatro pasos (Esquema 7). En el primer paso, la acetilación de **Int-2** procedió espontáneamente al agregar el anhídrido acético, esta tendencia fue general cuando se usaron derivados de anilina no desactivados por influencia de un grupo nitro en posición *orto* o *para*. Posteriormente, la nitración de **Int-3** llevó a la 4,5-dicloro-2-nitrofenilacetamida (**Int-4**) como producto principal, los productos secundarios se removieron lavando el producto crudo con metanol. En el tercer paso, **Int-4** se hidrolizó en medio ácido y generó **Int-5**, que mediante su reducción generó el 1,2-diaminobenceno correspondiente (**3**) con un buen rendimiento global (78%). Es importante resaltar que la protección del grupo amino juega un papel importante para dirigir la nitración a la posición 6 en **Int-3**, de otra manera la amina se protonaría en el medio ácido y la nitración se dirigiría preferentemente a la posición 5.

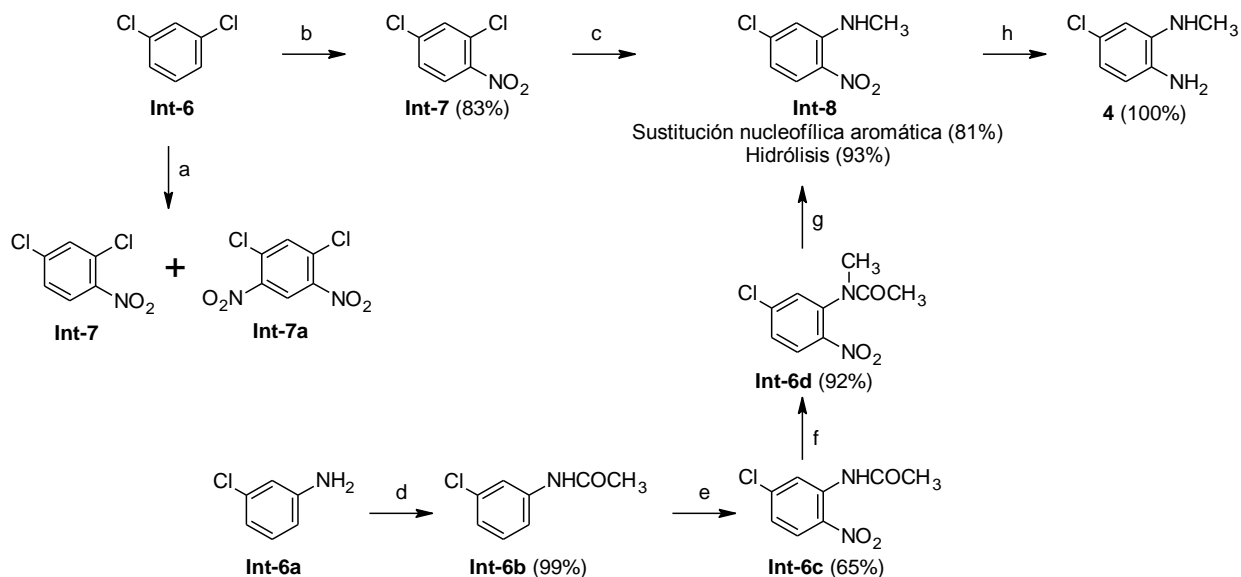


(a) Ac₂O; (b) HNO₃, H₂SO₄, 0 a 10 °C; (c) H₂SO₄, H₂O (1.2 eq.), Δ; (d) H₂, Ni-Raney, MeOH-AcOEt (50:50).

Esquema 7. Síntesis del compuesto 3.

7.4.2.3. 4-Cloro-N²-metil-1,2-diaminobenceno (**4**)

La preparación del compuesto **4** se realizó partiendo del 1,3-diclorobenceno comercial (**Int-6**) mediante la serie de reacciones mostradas en el Esquema 8.



(a) H₂SO₄, HNO₃ (60%), -5 a 5 °C; (b) H₂SO₄, HNO₃ (30%), -5 a 5 °C; (c) CH₃NH₂·HCl, K₂CO₃, DMSO/H₂O (80:20), 60 °C; (d) Ac₂O; (e) HNO₃ fum., Ac₂O; (f) NaOH, (CH₃)₂SO₄, monoglima, 30 a 40 °C; (g) H₂SO₄, H₂O (1.2 eq), 80 °C; (h) H₂, Ni-Raney, MeOH/AcOEt (50:50).

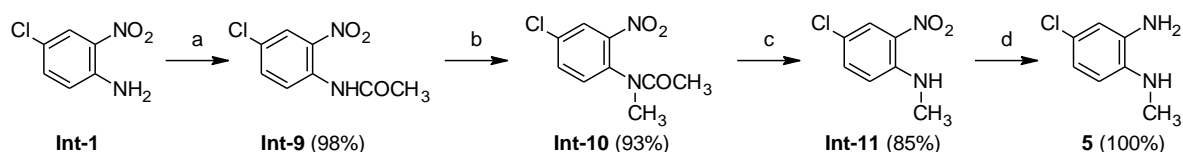
Esquema 8. Síntesis del compuesto **4**.

En la primera etapa se probó la reacción con la materia prima (**Int-6**) disuelta en ácido sulfúrico concentrado y se agregó ácido nítrico al 60%; sin embargo, se obtuvieron dos productos que se aislaron y correspondieron a **Int-7** con pf de 30.0–31.5 °C (lit: 31.5–33.0 °C) [95] y a **Int-7a** con pf de 101.1–102.8 °C (lit: 101–102 °C) [96]; no fue posible llevar a cabo la reacción a menor temperatura ya que el producto precipita. Una vez que se exploraron otras condiciones de nitración, los mejores resultados se obtuvieron usando ácido sulfúrico y ácido nítrico diluido (al 30%) manteniendo la temperatura entre -5 y 5 °C para dar una mezcla que contiene a **Int-7** como producto mayoritario, el cual se purificó por recristalización en hexano. Posteriormente, el intermediario **Int-7** se sometió a una sustitución nucleofílica; esta reacción se probó bajo distintas condiciones (distintos disolventes y temperaturas) dando en todos los casos productos no deseados en cantidades importantes, probablemente productos de sustitución en posición 4 o en ambas posiciones (2 y 4). Las mejores condiciones encontradas para generar una sustitución regioselectiva en

posición 2 fueron empleando como disolvente DMSO-H₂O (80:20) a 60 °C; de esta manera se obtuvo una cantidad mínima de productos secundarios, los cuales se separaron fácilmente del producto deseado por recristalización en metanol. Se probó un segundo método de preparación de **Int-8** partiendo de la 3-cloroanilina (**Int-6a**), esta ruta involucra una acetilación seguida de una nitración en ácido nítrico rojo fumante, una metilación y una hidrólisis. Además de que la última ruta involucra dos pasos más, el rendimiento global para obtener **Int-8** es un poco menor (55%) comparado con la primera ruta sintética (67%). Finalmente se obtuvo el compuesto **4**, con un rendimiento cuantitativo, mediante hidrogenación catalítica empleando Ni-Raney.

7.4.2.4. 4-Cloro-N¹-metil-1,2-diaminobenceno (**5**)

El compuesto **5** se preparó partiendo de la 2-nitro-4-cloroanilina comercial (**Int-1**) mediante una serie de cuatro reacciones (Esquema 9). La primera reacción requiere H₂SO₄ como catalizador dado a que el grupo nitro ejerce un efecto electroatractor y desactiva la amina para el ataque nucleofílico. Este comportamiento se observa para todos los derivados nitrados sometidos a acetilación con anhídrido acético estudiados en este trabajo. En el segundo paso, para la metilación de la amida presente en **Int-9** es necesario mantener la temperatura entre 30–35 °C para evitar la hidrólisis y la formación de productos secundarios. Posteriormente, **Int-10** se hidrolizó en ácido sulfúrico concentrado para generar la 4-cloro-N-metil-2-nitroanilina (**Int-11**) que por hidrogenación catalítica dio el compuesto **5**.



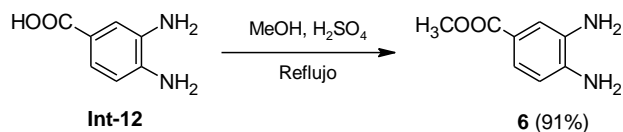
(a) Ac₂O, H₂SO₄ (cat.); (b) (CH₃)₂SO₄, NaOH, H₂O, monoglima; (c) H₂SO₄, H₂O (1.2 eq), Δ; (d) H₂, Ni-Raney, MeOH/AcOEt (50:50).

Esquema 9. Síntesis del compuesto **5**.

7.4.2.5. 3,4-Diaminobenzoato de metilo (**6**)

La preparación del compuesto **6** se realizó usando el ácido 3,4-diaminobenzoico comercial (**Int-12**) mediante una esterificación de Fischer empleando metanol y ácido

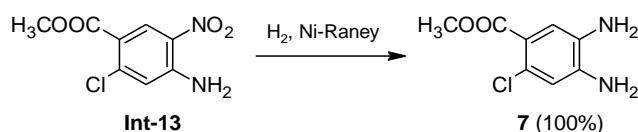
sulfúrico (Esquema 10). El producto mostró ser estable al aire a diferencia de muchos 1,2-diaminobencenos obtenidos en este trabajo. El punto de fusión para este compuesto fue de 104.1–106.1 °C, el cual coincidió con el reportado en la literatura (104 °C) [97].



Esquema 10. Síntesis del compuesto 6.

7.4.2.6. 4,5-Diamino-2-clorobenzoato de metilo (7)

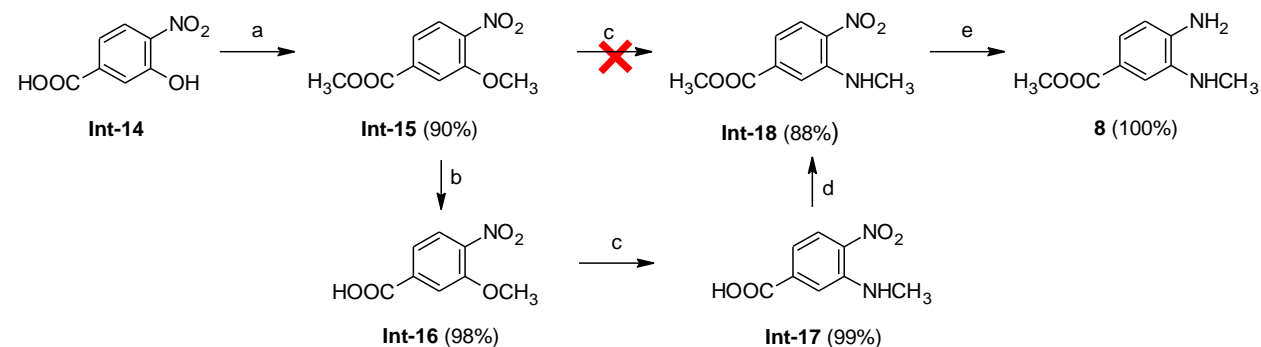
La síntesis del compuesto 7 se realizó por hidrogenación catalítica con Ni-Raney del 4-amino-2-cloro-5-nitrobenzoato de metilo (Int-13) que estaba disponible en nuestro laboratorio (Esquema 11).



Esquema 11. Síntesis del compuesto 7.

7.4.2.7. 4-Amino-3-(metilamino)benzoato de metilo (8)

Para la preparación del compuesto 8 se partió del ácido 3-hidroxi-4-nitrobenzoico comercial (Esquema 12).



(a) $(\text{CH}_3)_2\text{SO}_4$, K_2CO_3 , acetona, H_2O , 40 °C; (b) KOH , MeOH ; (c) $\text{CH}_3\text{NH}_2 \cdot \text{HCl}$, H_2O , K_2CO_3 , 105 °C; (d) MeOH , H_2SO_4 , Δ ; (e) H_2 , Pd/C 5% MeOH/AcOEt (50:50).

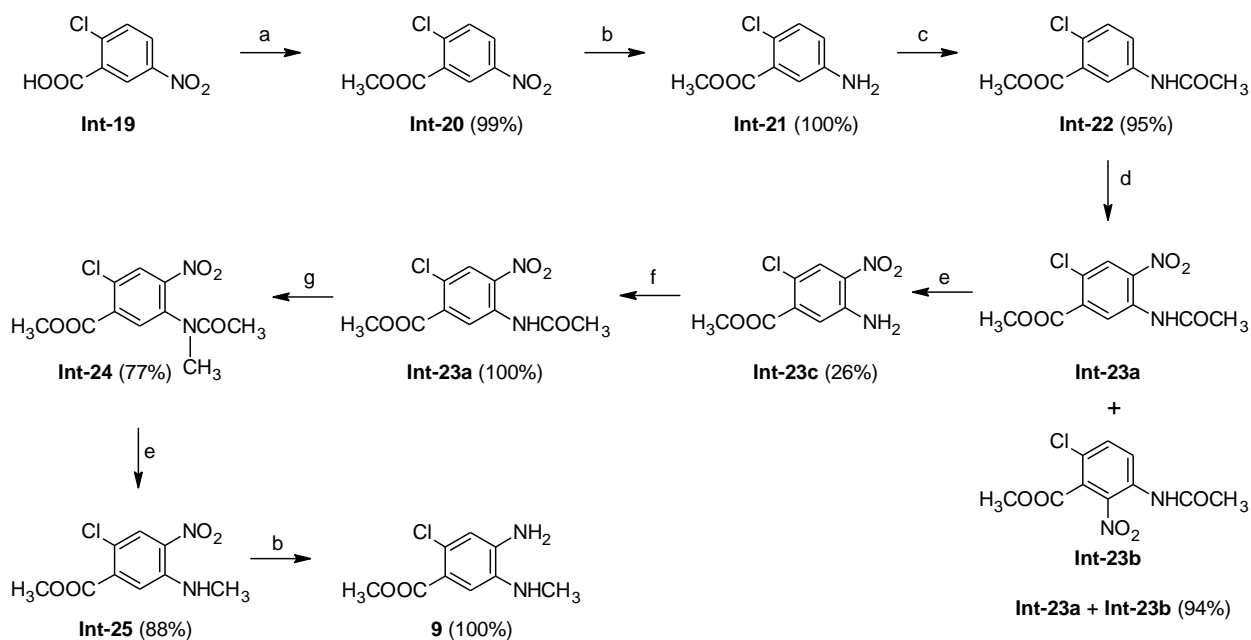
Esquema 12. Síntesis del compuesto 8.

Primero, **Int-14** se trató con sulfato de dimetilo en presencia de carbonato de potasio para dar **Int-15**. Posteriormente, **Int-15** se hidrolizó en condiciones básicas para dar lugar al ácido carboxílico **Int-16**; mismo que después se sometió a una reacción de sustitución nucleofílica aromática con metilamina para dar **Int-17**. El compuesto **Int-18** se preparó por esterificación de Fischer de **Int-17**. Por otro lado, la sustitución nucleofílica directa de **Int-15** para dar **Int-18** no fue una opción viable debido a que se obtiene una mezcla de productos donde el producto no deseado posiblemente es la amida correspondiente. Finalmente, por medio de una hidrogenación catalítica de **Int-18**, usando Pd/C, se llegó al producto deseado (**8**).

7.4.2.8. 4-Amino-2-cloro-5-(metilamino)benzoato de metilo (**9**)

Para la síntesis del compuesto **9** se usó el ácido 2-cloro-5-nitrobenzoico comercial (Esquema 13). El compuesto **Int-19** se trató con sulfato de dimetilo para formar el éster correspondiente (**Int-20**); posteriormente, mediante la hidrogenación catalítica de **Int-20** con Ni-Raney se obtuvo **Int-21** el cual se acetiló para generar **Int-22**. El siguiente paso fue la nitración del **Int-22** con ácido nítrico y sulfúrico; sin embargo, se obtuvo una mezcla de productos (**Int-23a** y **Int-23b**). Los productos **Int-23a** y **Int-23b** se caracterizaron por RMN. Para **Int-23a** se observaron dos singuletes en la región aromática (δ : 8.287 y 9.210). Por otra parte, para **Int-23b** se observó un doble de dobles (δ : 7.638, $J = 9.2$ Hz, 0.45 Hz) en la región aromática y un doblete (δ : 8.655, $J = 9$ Hz). La reacción se probó bajo distintas condiciones de nitración obteniendo en todos los casos resultados similares. La mezcla con los dos productos de nitración se hidrolizó en medio ácido y se logró separar selectivamente el compuesto **Int-23c** debido a que es insoluble en metanol. Posteriormente se acetiló **Int-23c** dando nuevamente **Int-23a** en forma pura. Una vez obtenido **Int-23a** se sometió a una metilación con sulfato de dimetilo para dar el producto **Int-24** que por hidrólisis ácida dio como producto el compuesto **Int-25**. Finalmente, el compuesto **9** se obtuvo por hidrogenación catalítica de **Int-25** usando Ni-Raney. El rendimiento global para esta ruta sintética, desde el **Int-19** hasta **9**, fue bajo (16%) debido principalmente a la etapa de nitración donde se forman dos isómeros difíciles de separar. Dado al impedimento estérico en la posición 6 del 5-acetamino-2-clorobenzoato de metilo (**Int-**

22) se esperaría que la posición 4 fuera la más favorecida a la sustitución electrofílica, sin embargo, los productos de nitración en 4 y 6 se obtienen en proporciones similares. Hasta el momento no se conoce el efecto de los sustituyentes en la orientación de la sustitución electrofílica hacia posición 6, no obstante, el carbonilo del grupo éster posiblemente juega un papel clave.

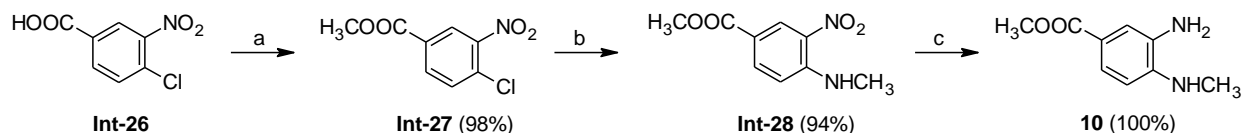


(a) $(\text{CH}_3)_2\text{SO}_4$, NaHCO_3 , DMF, 65°C , 2h; (b) H_2 , Ni-Raney, MeOH/AcOEt (50:50); (c) Ac_2O ; (d) H_2SO_4 , HNO_3 , 0 a 10°C ; (e) MeOH, H_2SO_4 , H_2O (1.2 eq), Δ ; (f) Ac_2O , H_2SO_4 (cat.); (g) $(\text{CH}_3)_2\text{SO}_4$, NaOH, H_2O , monoglisma.

Esquema 13. Síntesis del compuesto **9**.

7.4.2.9. 3-Amino-4-(metilamino)benzoato de metilo (**10**)

La preparación del compuesto **10** se llevó a cabo partiendo del ácido 4-cloro-3-nitrobenzoico comercial mediante una secuencia de tres pasos que se muestran en el Esquema 14. Primero se esterificó el ácido 4-cloro-3-nitrobenzoico comercial (**Int-26**) para dar **Int-27**. Posteriormente, mediante una sustitución nucleofílica con metilamina seguida de hidrogenación catalítica (Pd/C) se llegó de manera muy eficiente al producto **10**.

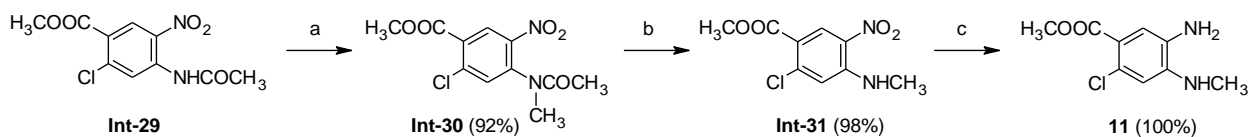


(a) $(\text{CH}_3)_2\text{SO}_4$, NaHCO_3 , DMF, 65 °C; (b) $\text{CH}_3\text{NH}_2 \cdot \text{HCl}$, DMF/ H_2O , K_2CO_3 , 130 °C; (c) H_2 , Pd/C, AcOEt/MeOH (50:50).

Esquema 14. Síntesis del compuesto **10**.

7.4.2.10. 5-Amino-2-cloro-4-(metilamino)benzoato de metilo (**11**)

El compuesto **11** se preparó a partir de la materia prima **Int-29**, que estaba disponible en nuestro laboratorio, mediante una secuencia de tres pasos (Esquema 15). Las reacciones empleadas incluyen procesos previamente descritos de metilación para dar **Int-30**, hidrólisis ácida para dar **Int-31** e hidrogenación catalítica usando Ni-Raney para obtener **11**. Todos los intermediarios se obtuvieron en rendimientos excelentes.

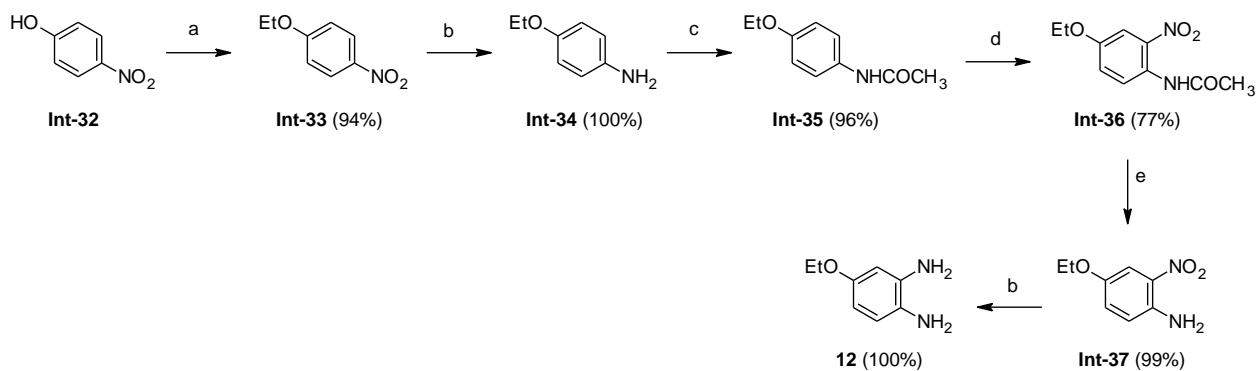


(a) $(\text{CH}_3)_2\text{SO}_4$, NaOH, H_2O , monoglisma; (b) MeOH, H_2SO_4 , H_2O (1.2 eq) reflujo; (c) H_2 , Ni-Raney, MeOH/AcOEt (50:50).

Esquema 15. Síntesis del compuesto **11**.

7.4.2.11. 1,2-Diamino-4-etoxibenceno (**12**)

El compuesto **12** se preparó mediante una secuencia de seis reacciones (Esquema 16). Primero, el 4-nitrofenol comercial (**Int-32**) se alquiló con yodoetano para dar **Int-33**, que por hidrogenación catalítica dio **Int-34**. Posteriormente, mediante una reacción de acetilación con anhídrido acético se protegió el grupo amino (**Int-35**) para después realizar una nitración que dio una mezcla de dos isómeros; sin embargo, el producto mayoritario (**Int-36**) se purificó fácilmente por recristalización con etanol. Finalmente, mediante hidrólisis básica de **Int-36** y una posterior reducción del grupo nitro se llegó al compuesto **12**. Todos los pasos en esta ruta se llevaron a cabo en rendimientos excelentes con excepción del producto de nitración **Int-36** que se obtuvo en un 77%.

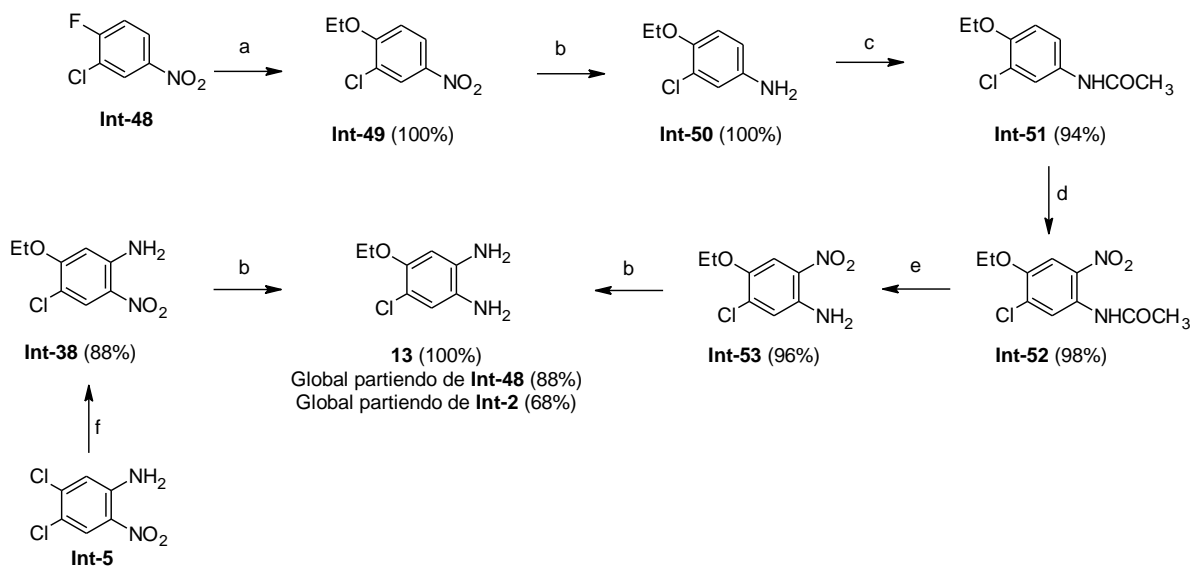


(a) $\text{CH}_3\text{CH}_2\text{I}$, K_2CO_3 , DMF/ H_2O , Δ ; (b) H_2 , Pd/C, MeOH/AcOEt (50:50); (c) Ac_2O ; (d) HNO_3 , AcOH/ H_2O , 10 a 40 °C; (e) KOH, MeOH.

Esquema 16. Síntesis del compuesto **12**.

7.4.2.12. 1,2-Diamino-4-cloro-5-etoxibenceno (**13**)

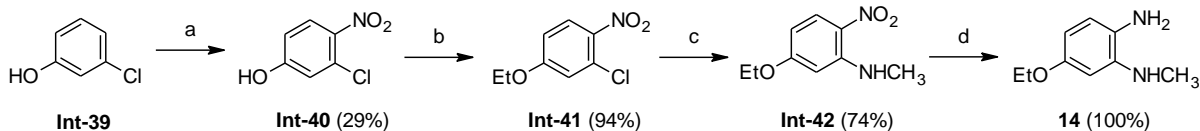
La preparación del compuesto **13** se realizó por dos rutas sintéticas diferentes (Esquema 17). La primera de ellas parte de 3-cloro-4-fluoronitrobenzenu comercial (**Int-48**), el cual se sometió a una reacción de sustitución nucleofílica con etóxido de sodio (preparado previamente con sodio metálico y etanol anhidro) para dar **Int-49** que por hidrogenación con Ni-Raney dio **Int-50**. Éste último se acetiló para dar **Int-51** que se sometió a una nitración dando exclusivamente el compuesto **Int-52**. La hidrólisis de **Int-52** en medio básico dio la 2-nitroanilina **Int-53** que por hidrogenación catalítica dio el producto **13**. Se exploró un segundo método para preparar **13** a partir de **Int-5** (previamente preparado, Esquema 7). La sustitución nucleofílica sobre **Int-5** con etóxido de sodio y posterior reducción por hidrogenación llevó a **13**. Al comparar ambas rutas se observó que aunque la síntesis a partir de **Int-48** es más larga (seis pasos), el rendimiento global es superior (88%). Por otra parte, la ruta alterna permitió preparar **13** en dos pasos con un buen rendimiento cuando se cuenta con **Int-5** (88%). Sin embargo, preparar **Int-5** a partir de la 3,4-dicloroanilina (**Int-2**) comercial implica tres pasos de reacción adicionales bajando el rendimiento global (68%) y haciéndolo inferior al de la primera ruta (88%).



Esquema 17. Síntesis del compuesto **13**.

7.4.2.13. 1,2-Diamino-4-etoxi-*N*²-metilbenceno (**14**)

La preparación del compuesto **14** se realizó partiendo del 3-clorofenol (**Int-39**) como se muestra en el Esquema 18. El primer paso fue una nitración para obtener una mezcla de tres productos: el 3-cloro-4-nitrofenol (**Int-40**) en proporción mayoritaria y dos productos secundarios no caracterizados (probablemente el 5-cloro-2-nitrofenol y el 3-cloro-2-nitrofenol) [98]. La mezcla de productos de nitración se recrystalizó de benceno para dar 29% de **Int-40** en forma pura. Los siguientes pasos consistieron en la formación del éter correspondiente (**Int-41**), seguido de una sustitución nucleofílica con metilamina para dar **Int-42**. Finalmente, la hidrogenación catalítica de **Int-42** dio el producto **14**. El rendimiento global para esta ruta fue bajo (20%), en gran medida por el paso de nitración.

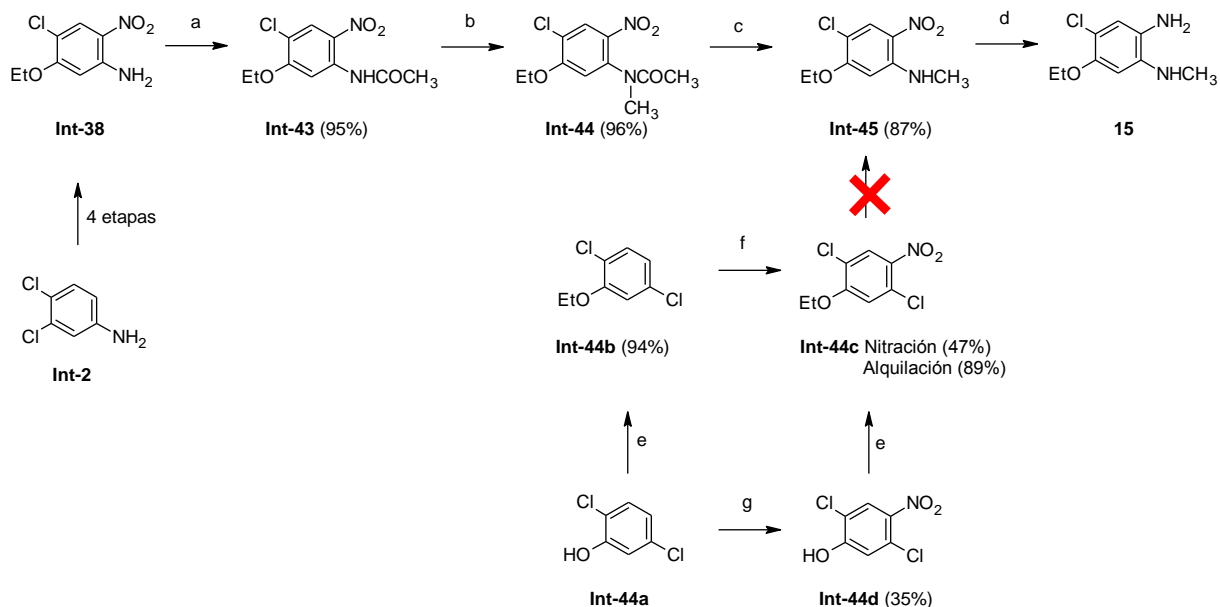


(a) AcOH , HNO_3 , -5 a 0 °C; (b) $\text{CH}_3\text{CH}_2\text{I}$, K_2CO_3 , $\text{DMF/H}_2\text{O}$, Δ ; (c) $\text{CH}_3\text{NH}_2\cdot\text{HCl}$, $\text{DMF/H}_2\text{O}$, K_2CO_3 , 110 °C; (d) H_2 , Pd/C (5%), AcOEt-MeOH (50:50).

Esquema 18. Síntesis del compuesto **14**.

7.4.2.14. 1,2-Diamino-4-cloro-5-etoxi-*N*¹-metilbenceno (**15**)

El compuesto **15** se preparó a partir de **Int-38** como se muestra en el Esquema 19.



(a) Ac_2O , H_2SO_4 (cat.); (b) $(\text{CH}_3)_2\text{SO}_4$, NaOH , H_2O , monoglíma; (c) H_2SO_4 , MeOH , H_2O (1.2 eq); (d) H_2 , Ni-Raney; (e) $\text{CH}_3\text{CH}_2\text{I}$, K_2CO_3 , $\text{DMF}/\text{H}_2\text{O}$, Δ ; (f) $\text{HNO}_3/\text{H}_2\text{SO}_4$, CHCl_3 , 25 a 30 °C; (g) HNO_3/AcOH , 10 a 15 °C.

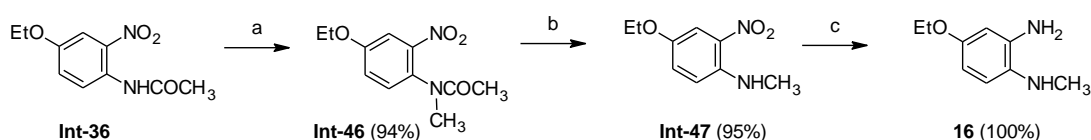
Esquema 19. Síntesis del compuesto **15**.

El compuesto **Int-38** se acetiló para dar la amida **Int-43** que posteriormente dio **Int-44** por metilación. La hidrólisis de **Int-44** dio la 2-nitroanilina correspondiente **Int-45** que por reducción con Ni-Raney dio el compuesto **15**. Todos los pasos en esta secuencia se llevaron a cabo con excelentes rendimientos. Por otro lado se probaron rutas alternas a partir del 2,5-diclorofenol comercial (**Int-44a**) con la finalidad de obtener el compuesto **15** en un menor número de pasos ya que la síntesis total de **15** a partir de **Int-2** involucra ocho etapas. La primera ruta consiste en la alquilación de **Int-44a** con yodoetano para dar **Int-44b** que por una reacción de nitración dio **Int-44c**. Esta ruta no fue muy conveniente debido a que el producto de nitración se obtiene en un rendimiento bajo por la formación de productos secundarios. Se probaron varias condiciones de nitración siendo la reacción con mezcla sulfonítrica en cloroformo la mejor de ellas. Por otra parte se probó la nitración de **Int-44a** y dio el producto **Int-44d** en rendimiento bajo, el cual se alquiló para dar **Int-44c**. El tercer paso en ambas rutas

consistió en la sustitución nucleofílica para dar **Int-45**; sin embargo, se obtuvo una mezcla dos productos en proporciones 50:50, lo que sugiere que además del cloro en posición *orto* el grupo etoxilo también actúa como grupo saliente para dirigir la sustitución en posición *para*. Por esta razón, aunque involucre un mayor número de pasos, es más conveniente la preparación de **15** a partir de **Int-2**.

7.4.2.15. 1,2-Diamino-4-etoxi-*N*¹-metilbenceno (**16**)

La preparación del compuesto **16** se realizó a partir del **Int-36** preparado previamente (Esquema 20). El primer paso consistió en la metilación de **Int-36** para dar **Int-46**, que por una hidrólisis en medio básico dio **Int-47**. Finalmente, mediante hidrogenación catalítica de **Int-47** se obtuvo el compuesto **16**.

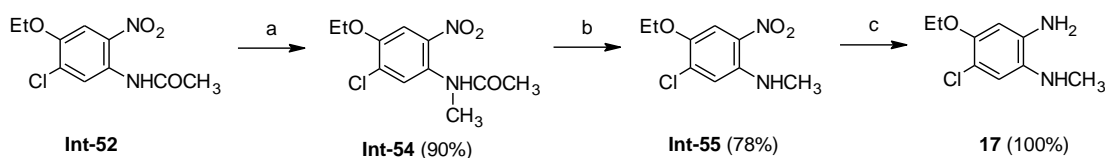


(a) $(\text{CH}_3)_2\text{SO}_4$, NaOH/H₂O, monoglíma; (b) KOH, MeOH, Δ ; (c) H₂, Pd/C (5%), MeOH/AcOEt.

Esquema 20. Síntesis del compuesto **16**.

7.4.2.16. 1,2-Diamino-5-cloro-4-etoxi-*N*¹-metilbenceno (**17**)

El compuesto **17** se preparó a partir del **Int-52** mediante tres etapas que incluyen metilación de la amida para dar **Int-54**, hidrólisis básica para dar **Int-55** y, finalmente, hidrogenación catalítica con Ni-Raney para dar el producto deseado (**17**) (Esquema 21).



(a) $(\text{CH}_3)_2\text{SO}_4$, NaOH/H₂O, monoglíma; (b) KOH, MeOH, Δ ; (c) H₂, Ni-Raney, MeOH/AcOEt (50:50).

Esquema 21. Síntesis del compuesto **17**.

7.4.3. Síntesis de los 2-mercaptobencimidazoles (18–34)

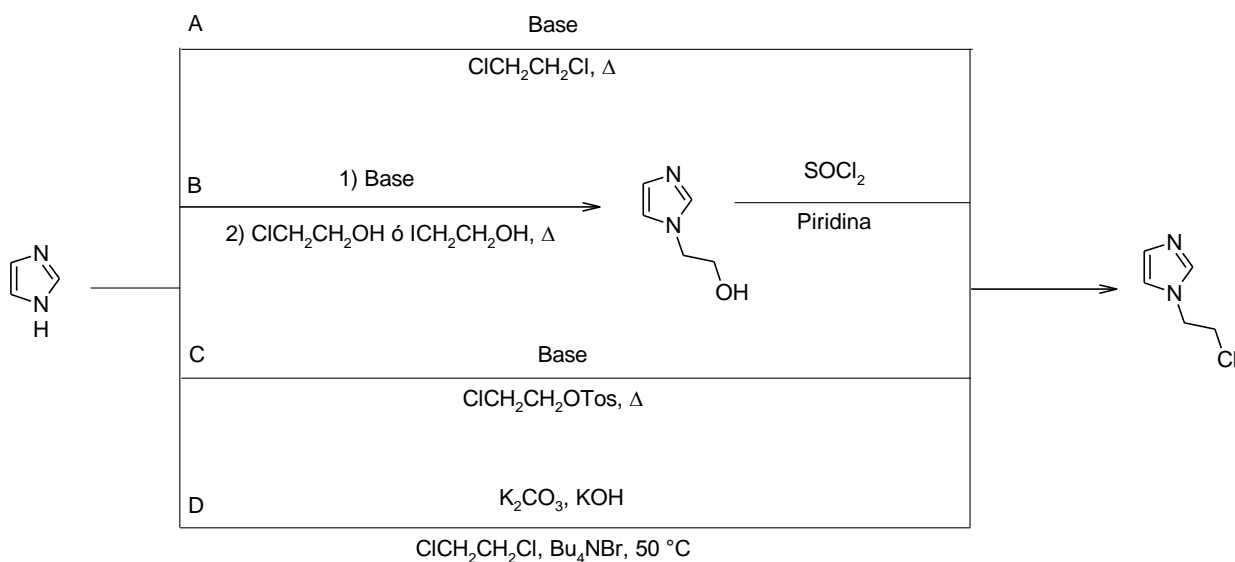
Todos los intermediarios del tipo 2-mercaptobencimidazol (18–34) se prepararon empleando disulfuro de carbono, hidróxido de potasio, etanol-agua y el 1,2-diaminobenceno apropiadamente sustituido (1–17) [70]. Los rendimientos y los puntos de fusión de los 2-mercaptobencimidazoles sintetizados se muestran en la Tabla 20. Los compuestos se obtuvieron en rendimientos que van de muy buenos a pobres (86–42%). El rendimiento más bajo (42%) corresponde al compuesto **26** que se preparó a partir del 1,2-diaminobenceno **9**. Sólo en este caso la reacción se llevó a cabo en el doble del tiempo esperado y requirió el doble de equivalentes de disulfuro de carbono. Por último, en todos los casos, fue necesario recrystalizar los productos y tratarlos con carbón activado, debido a que se forman compuestos coloridos por la presencia de productos secundarios resultantes de la oxidación de los 1,2-diaminobencenos.

Tabla 20. Rendimiento y punto de fusión de los 2-mercaptobencimidazoles sintetizados

Clave	R ₁	R ₂	R ₃	% Rend.	Pf (°C)	Pf [lit] (°C)
18	H	H	H	85	301.0–304.6	303–304[70]
19	H	Cl	H	56	327.0–329.0	327–329[11]
20	H	Cl	Cl	73	344.0–346.0	344–346[11, 99]
21	CH ₃	H	Cl	67	280.0–282.5	
22	CH ₃	Cl	H	50	233.0–234.3	230–231[100]
23	H	COOCH ₃	H	62	306.1–309.0	
24	H	COOCH ₃	Cl	57	243.9–244.8	
25	CH ₃	H	COOCH ₃	56	243.2–243.8	
26	CH ₃	Cl	COOCH ₃	42	230.5–231.4	
27	CH ₃	COOCH ₃	H	77	251.4–252.3	
28	CH ₃	COOCH ₃	Cl	65	250.2–251.8	
29	H	OCH ₂ CH ₃	H	86	246.5–248.5	245[101]
30	H	OCH ₂ CH ₃	Cl	78	300.0–300.1	
31	CH ₃	H	OCH ₂ CH ₃	85	179.0–181.3	
32	CH ₃	Cl	OCH ₂ CH ₃	80	174.7–175.8	
33	CH ₃	OCH ₂ CH ₃	H	81	151.7–153.4	
34	CH ₃	OCH ₂ CH ₃	Cl	69	248.5–250.6	

7.4.4. Síntesis del 1-(2-Cloroetil)-1H-imidazol

La síntesis del intermediario 1-(2-cloroetil)-1H-imidazol se abordó usando varias estrategias de alquilación del imidazol. Algunos de los métodos planteados se muestran de manera general en el Esquema 22. La alquilación del imidazol, por las vías A–C, se realizó utilizando distintas bases como DBU, Et₃N, K₂CO₃ y NaH; también se probó la reacción en distintos disolventes como DMF, CH₃CN, acetona e incluso se probó el agente alquilante como disolvente (ClCH₂CH₂Cl ó HOCH₂CH₂Cl). En la mayoría de los casos, el imidazol y el producto se observaron en cromatografía en capa fina en proporciones aproximadamente de 50:50, respectivamente, sin lograr un mayor avance de la reacción al transcurrir el tiempo. Por otro lado, cuando se usó en exceso el agente alquilante, el producto de interés y la materia prima se consumieron formando un producto muy polar soluble en agua que probablemente sea un producto iónico derivado de la doble alquilación del imidazol. Adicionalmente, se probaron las reacciones usando 1,2-dicloroetano o el tosilato derivado del 2-cloroetano (2-cloroetil-4-metilbencensulfonato) y la sal sódica del imidazol obteniendo resultados similares. Cuando se intentó recuperar el producto antes de la desaparición total de la materia prima se obtuvieron en todos los casos rendimientos muy pobres.



Esquema 22. Rutas planteadas para la preparación del 1-(2-Cloroetil)-1H-imidazol.

El método más conveniente para la preparación del 1-(2-cloroetil)-1*H*-imidazol fue mediante la técnica reportada previamente por Bogdal y Jaskot (ruta D, Esquema 22) [71]. Este método consiste en una reacción sólido-líquido y hace uso de un catalizador de transferencia de fases. Esta metodología tiene como inconveniente que genera en menor proporción el producto de eliminación (vinilimidazol), sin embargo, es la metodología que proporciona un mayor rendimiento del producto deseado.

7.4.5. Síntesis de los compuestos (**S1J01–S3J19**)

Partiendo de los derivados de 2-mercaptobencimidazol **18–34** se obtuvieron los compuestos finales **S1J01–S1J03**, **S1J05–S1J06** y **S2J08–S3J19** (Tabla 21). Primero se preparó el anión del 2-mercaptobencimidazol correspondiente (**18–22** y **29–34**) usando KOH al 50%; para los compuestos que presentan un grupo éster (**23–28**) se usó DBU para evitar la hidrólisis. Posteriormente, se realizó una reacción de sustitución nucleofílica sobre el 1-(2-cloroetil)-1*H*-imidazol. Los rendimientos van de moderados a buenos (52–90%, Tabla 21), ya que se obtiene como producto secundario el derivado de la alquilación sobre el nitrógeno del bencimidazol (ver Figura 41). Los compuestos **S1J04** y **S1J07** se obtuvieron mediante *N*-metilación de los compuestos **S1J01** y **S1J03** respectivamente. Aunque el rendimiento de esta última reacción fue muy bueno (84–86%), se percibió por cromatografía en capa fina un producto secundario (o productos secundarios) muy polar que aunque no fue caracterizado, se sugiere que se trata de las sales producidas por la alquilación del azufre y de los nitrógenos sobre la posición 3 del bencimidazol e imidazol.

Tabla 21. Rendimiento y punto de fusión para los compuestos finales

A)

18-34

1) KOH o DBU, CH₃CN
2) 1-(2-Cloroetil)-1H-imidazol, 60 °C, 6-8h

S1J01–S1J03
S1J05–S1J06
S2J08–S3J19

B)

S1J01
S1J03

1) KOH, (CH₃)₂CO
2) CH₃I, 25 °C, 4h

S1J04
S1J07

Clave	R ₁	R ₂	R ₃	% Rend.	Pf (°C)
S1J01	H	H	H	55	139.0–140.1
S1J02	H	Cl	H	87	110.8–112.5
S1J03	H	Cl	Cl	77	197.8–199.9
S1J04	CH ₃	H	H	86	51.5–52.7
S1J05	CH ₃	H	Cl	81	119.1–120.5
S1J06	CH ₃	Cl	H	71	110.2–111.7
S1J07	CH ₃	Cl	Cl	84	142.5–144.0
S2J08	H	COOCH ₃	H	85	182.3–183.5
S2J09	H	COOCH ₃	Cl	69	197.5–198.9
S2J10	CH ₃	H	COOCH ₃	60	119.6–120.5
S2J11	CH ₃	Cl	COOCH ₃	63	136.8–138.8
S2J12	CH ₃	COOCH ₃	H	52	116.3–117.9
S2J13	CH ₃	COOCH ₃	Cl	70	126.9–128.3
S3J14	H	OCH ₂ CH ₃	H	80	96–105
S3J15	H	OCH ₂ CH ₃	Cl	82	123.4–126.0
S3J16	CH ₃	H	OCH ₂ CH ₃	90	53.0–57.1
S3J17	CH ₃	Cl	OCH ₂ CH ₃	75	124.4–126.5
S3J18	CH ₃	OCH ₂ CH ₃	H	72	111.1–113.3
S3J19	CH ₃	OCH ₂ CH ₃	Cl	70	139.7–142.2

7.4.6. Caracterización de compuestos finales

Los compuestos obtenidos (**S1J01–S3J19**) se caracterizaron utilizando técnicas espectroscópicas y espectrométricas. En las siguientes secciones se muestra el análisis de los espectros de resonancia magnética nuclear y de masas.

7.4.6.1. RMN ¹H de los compuestos finales

En los espectros de RMN ¹H se observan características comunes para todos los compuestos debido a que contienen la misma estructura base. En la Figura 40 se muestra un resumen esquemático de las señales observadas en los espectros de RMN ¹H y su relación con la estructura general de las moléculas finales.

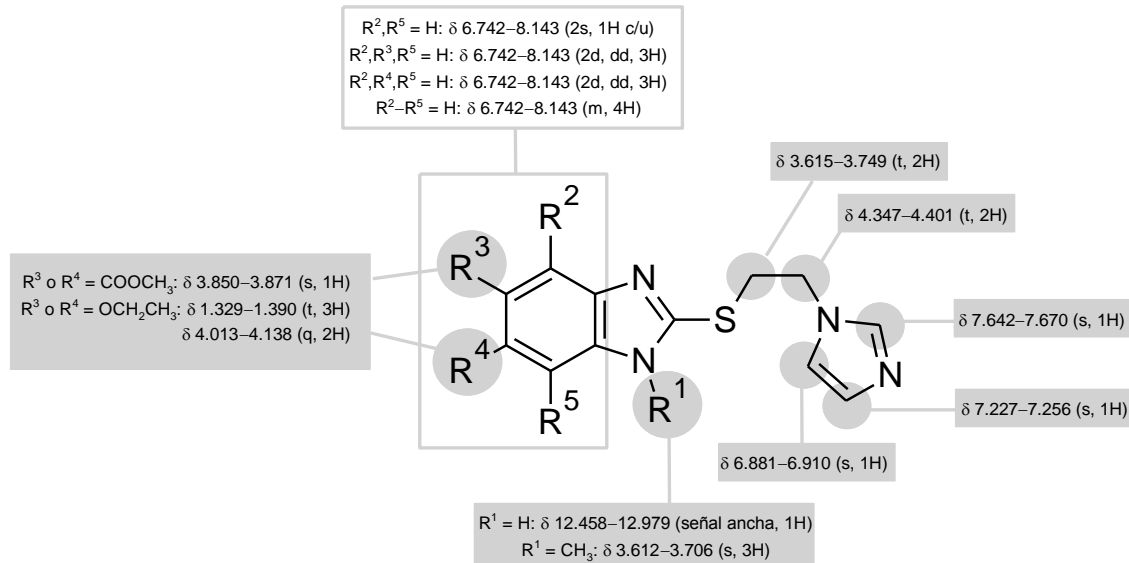


Figura 40. Relación entre la estructura general de los compuestos finales y las señales de RMN ¹H observadas.

Adicionalmente, en las Tablas 22 y 23 se muestran los desplazamientos químicos, la multiplicidad y las constantes de acoplamiento para los protones en cada compuesto. De manera general, para todos los compuestos se observan dos tripletes correspondientes a dos metilenos que unen el derivado de 2-mercaptobencimidazol y el imidazol. El primer triplete se observa para todos los compuestos en un rango de 3.615–3.749 y corresponde al metileno unido al azufre; el segundo, se observa a 4.347–4.401 ppm y corresponde al metileno unido al nitrógeno del anillo de imidazol (Tabla 22).

Algunas señales solamente son comunes en algunos casos según el patrón de sustitución. Cuando los compuestos se encuentran sustituidos en posición 1 por hidrógeno o metilo se observa una señal ancha característica a 12.458–12.979 ppm (para NH, Tabla 23) o un singulete a 3.612–3.706 ppm (para NCH₃), respectivamente. Por otra parte, para los derivados de la serie 2, cuando se tiene un éster en posiciones

5 ó 6 se observa un singulete a 3.850–3.871 ppm y para los derivados de la serie 3 encontramos un triplete a 1.329–1.390 ppm que corresponde al metilo del grupo etoxilo y un cuadruplete a 4.013–4.138 ppm que corresponde al metileno del grupo etoxilo.

Tabla 22. Resumen de las señales de RMN ¹H encontradas a campo alto para los compuestos finales

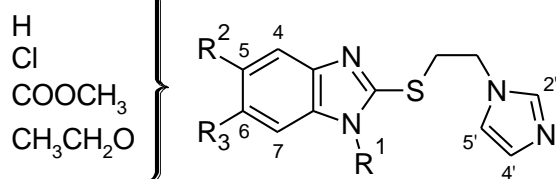
No.						
	NCH ₃	SCH ₂ CH ₂ N	SCH ₂ CH ₂ N	COOCH ₃	OCH ₂ CH ₃	OCH ₂ CH ₃
S1J01		δ: 3.664, t, J = 6.45 Hz	δ: 4.371, t, J = 6.45 Hz			
S1J02		δ: 3.667, t, J = 6.3 Hz	δ: 4.370, t, J = 6.3 Hz			
S1J03		3.665, t, J = 6.6 Hz	4.364, t, J = 6.6 Hz			
S1J04	3.653, s	δ: 3.708, t, J = 6.3 Hz	δ: 4.385, t, J = 6.45 Hz			
S1J05	3.639, s	δ: 3.706, t, J = 6.45 Hz	4.381, t, J = 6.45 Hz			
S1J06	3.653, s	δ: 3.712, t, J = 6.45 Hz	4.376, t, J = 6.45 Hz			
S1J07	3.633, s	δ: 3.709, t, J = 6.45 Hz	4.376, t, J = 6.3 Hz			
S2J08		δ: 3.702, t, J = 6.45 Hz	4.395, t, J = 6.45 Hz	3.851, s		
S2J09		δ: 3.690, t, J = 6.45 Hz	4.381, t, J = 6.45 Hz	3.850, s		
S2J10	3.706, s	δ: 3.749, t, J = 6.45 Hz	4.401, t, J = 6.45 Hz	3.871, s		
S2J11	3.685, s	3.741, t, J = 6.45 Hz	4.388, t, J = 6.45 Hz	3.871, s		
S2J12	3.688, s	3.737, t, J = 6.45 Hz	4.401, t, J = 6.3 Hz	3.862, s		
S2J13	3.662, s	3.722, t, J = 6.3 Hz	4.389, t, J = 6.3 Hz	3.856, s		
S3J14		3.615, t, J = 6.6 Hz	4.347, t, J = 6.6 Hz		4.013, q, J = 7.06 Hz	1.329, t, J = 7.2 Hz
S3J15		3.627, t, J = 6.45 Hz	4.351, t, J = 6.45 Hz		4.092, q, J = 7.0 Hz	1.367, t, J = 7.05 Hz
S3J16	3.612, s	3.645, t, J = 6.3 Hz	4.354, t, J = 6.3 Hz		4.055, q, J = 7.0 Hz	1.347, t, J = 7.1 Hz
S3J17	3.633, s	3.661, t, J = 6.3 Hz	4.356, t, J = 6.45 Hz		4.138, q, J = 7.00 Hz	1.390, t, J = 6.9 Hz
S3J18	3.616, s	3.674, t, J = 6.45 Hz	4.371, t, J = 6.45 Hz		4.026, q, J = 6.9 Hz	1.334, t, J = 6.9 Hz
S3J19	3.616, s	3.678, t, J = 6.45 Hz	4.374, J = 6.45 Hz		4.105, q, J = 6.9 Hz	1.370, t, J = 7.05 Hz

Otra de las características comunes para todos los compuestos finales son los hidrógenos ubicados sobre el anillo de imidazol que se observan como tres singuletes con intensidad similar; el primero de ellos se encuentra en un rango de desplazamiento de 6.881–6.910 ppm y corresponde al hidrógeno en posición 5 (H5'); el segundo, correspondiente al hidrógeno de la posición 4 (H4'), se encuentra a 7.227–7.256 ppm; y el tercero, correspondiente al hidrógeno de la posición 2 (H2'), se encuentra a 7.642–7.670 ppm (Tabla 23). El número y multiplicidad de las señales correspondientes a la parte bencenoide del bencimidazol son variables según el patrón de sustitución (Tabla 23). Los compuestos sin sustitución presentaron señales múltiples; para los compuestos sustituidos en posición 5 ó 6 se pueden distinguir tres señales correspondientes a dos dobletes y un doble de dobles (la multiplicidad varía en algunos casos respecto a la calidad del espectro). Los compuestos sustituidos en ambas posiciones presentan dos singuletes; con excepción del compuesto **S1J03** donde se observa una sola señal que integra para dos protones aromáticos a 7.723 ppm.

En la Figura 41A se muestra un ejemplo de un espectro de RMN ^1H correspondiente al compuesto **S2J10** donde se observan las señales características descritas previamente. Es interesante comparar este espectro con el obtenido para el producto secundario derivado de la reacción de alquilación con 1-(2-cloroetil)-1*H*-imidazol. En ambos espectros se pueden observar características muy similares en cuanto a multiplicidad y desplazamiento químico, sin embargo, en la Figura 41B las señales correspondientes al metileno unido a la posición 3 y al hidrógeno sobre la posición 4 del bencimidazol se desplazan significativamente con respecto al producto principal **S2J10**. Este resultado confirma que la alquilación se llevó en una posición distinta. Con base en el desplazamiento químico de las señales se propone que el producto secundario es el producto de alquilación sobre el nitrógeno del bencimidazol. Es de esperarse la formación de este producto en algún grado dado a que la forma de tiona es muy estable en los derivados de 2-mercaptobencimidazol.

Tabla 23. Resumen de las señales de RMN ¹H encontradas a campo bajo para los compuestos finales

No.	H4	H7	H5	H6	H2'	H4'	H5'	NH
S1J01	7.462, bs		7.129, m		7.654, s	7.242, s	6.910, s	12.654, bs
S1J02	7.527, s	7.455, d, J = 7.5 Hz		7.148, dd, J = 8.6; 1.4 Hz	7.662, s	7.242, s	6.908, s	12.858, bs
S1J03	7.723, s	7.723, s			7.642, s	7.227, s	6.889, s	12.979, bs
S1J04	7.564–7.593, m	7.469–7.499, m		7.143–7.231, m	7.660, s	7.248, s	6.903, s	
S1J05	7.557, d, J = 8.4 Hz	7.655, s	7.177, dd, J = 8.4; 2.1 Hz		7.661, s	7.240, s	6.896, s	
S1J06	7.630, d, J = 1.8 Hz	7.525, d, J = 8.7 Hz		7.224, dd, J = 8.7; 1.8 Hz	7.649, s	7.238, s	6.887, s	
S1J07	7.867, s	7.802, s			7.652, s	7.234, s	6.885, s	
S2J08	8.047, s	7.537, d, J = 8.4 Hz		7.785, dd, J = 8.55; 1.65 Hz	7.668, s	7.246, s	6.905, s	12.902, bs
S2J09	7.909, s	7.633, s			7.660, s	7.240, s	6.896, s	12.908, bs
S2J10	7.634, d, J = 8.4 Hz	8.094, d, J = 1.2 Hz	7.810, dd, J = 8.55; 1.65 Hz		7.667, s	7.252, s	6.895, s	
S2J11	7.733, s	8.011, s			7.667, s	7.243, s	6.881, s	
S2J12	8.143, d, J = 1.2 Hz	7.598, d, J = 8.4 Hz		7.842, dd, J = 8.55; 1.65 Hz	7.670, s	7.256, s	6.890, s	
S2J13	7.996, s	7.788, s			7.663, s	7.245, s	6.884, s	
S3J14	7.080, bs	7.397, bs		6.742, dd, J = 9.2; 2.8 Hz	7.649, s	7.233, s	6.903, s	12.458, bs
S3J15	7.071, bs	7.757, bs			7.647, s	7.230, s	6.899, s	12.644, bs
S3J16	7.439, d, J = 8.7 Hz	7.064, d, J = 2.4 Hz	6.772, dd, J = 8.85; 2.25 Hz		7.655, s	7.239, s	6.898, s	
S3J17	7.622, s	7.288, s			7.645, s	7.230, s	6.889, s	
S3J18	7.123, d, J = 2.4 Hz	7.352, d, J = 9.0 Hz		6.815, dd, J = 8.7; 2.4 Hz	7.661, s	7.244, s	6.907, s	
S3J19	7.314, s	7.660, s			7.660, s	7.240, s	6.899, s	



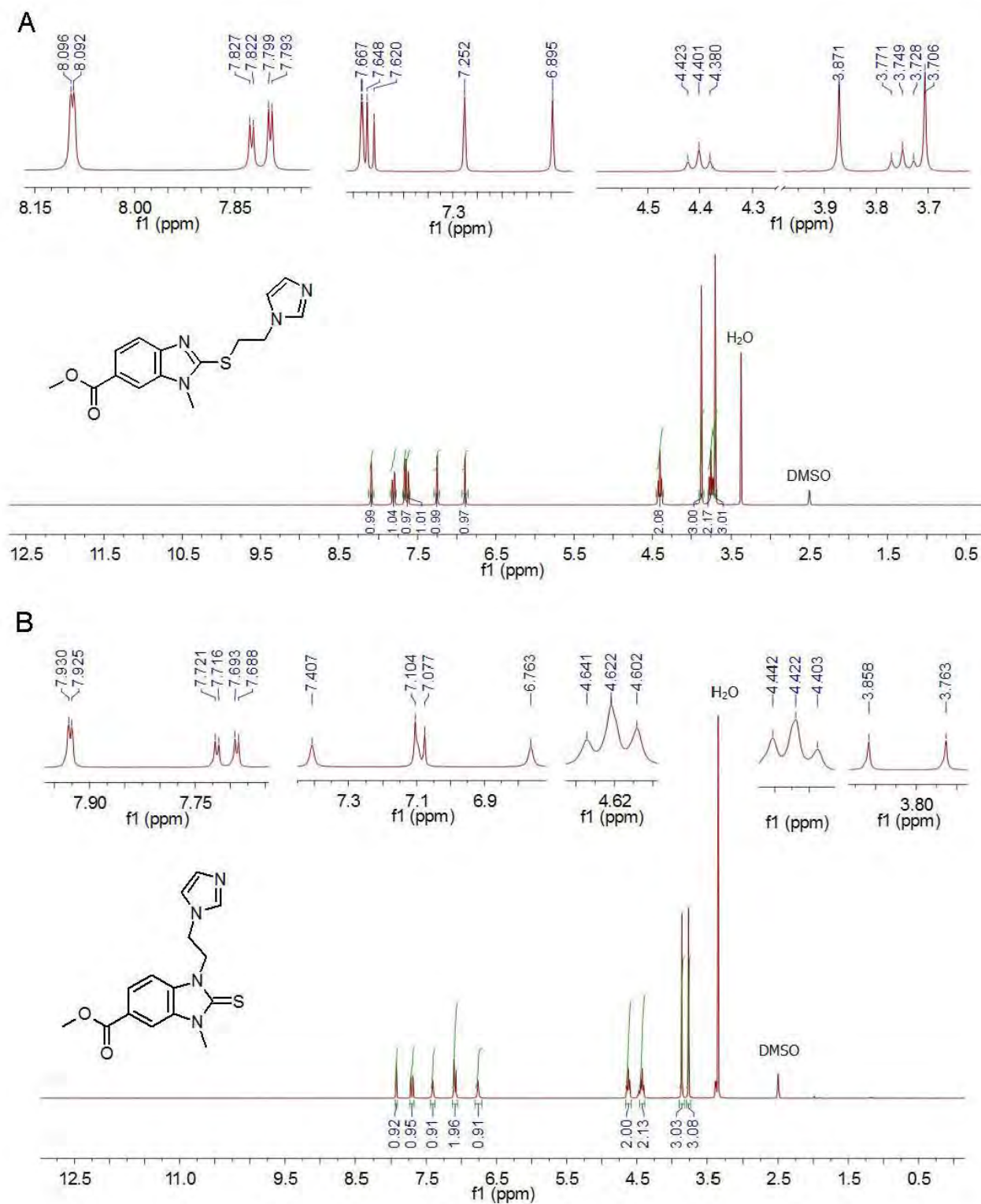


Figura 41. Espectro de RMN ^1H del compuesto **S2J10** (A) y producto secundario (B).

7.4.6.2. RMN ^{13}C de los compuestos finales

Los espectros de RMN ^{13}C presentan características comunes para todos los compuestos (Figura 42). Se observó la señal correspondiente al carbono alifático unido a azufre dentro de un rango de 31.99–33.17 ppm y el carbono unido al nitrógeno del imidazol aparece en un rango de 45.08–45.90 ppm (Tabla 24). Por otro lado, se observaron algunas características particulares para las series 2 y 3. Para la serie 2 aparece el carbono sp^3 del grupo éster en un rango de 51.97–52.39 y el carbono carbonílico a campo bajo entre 165.81–166.75. Para la serie 3 se observaron a campo alto dos señales correspondientes al metilo (14.61–15.21) y al metileno (63.42–65.43) del grupo etoxilo.

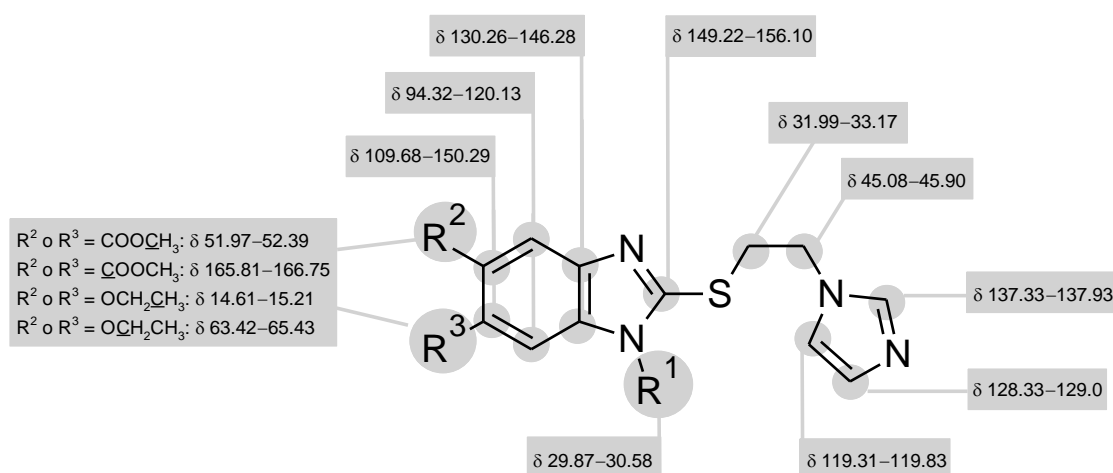


Figura 42. Relación entre la estructura general de los compuestos finales y las señales de RMN ^{13}C observadas.

Otras de las señales características para todos los compuestos corresponden a los carbonos situados en el anillo de imidazol. Los carbonos de las posiciones 2, 4 y 5 del anillo de imidazol se encuentran en un rango de 137.33–137.93, 128.33–129.00 y 119.31–119.83, respectivamente. Este orden en el desplazamiento químico es esperado debido a que el carbono en la posición 2 está unido directamente a dos nitrógenos que lo hacen deficiente en electrones. Por otra parte, aunque los carbonos del imidazol situados en posiciones 4 y 5 están unidos directamente a un nitrógeno, la diferencia en sus desplazamientos se puede explicar por el efecto inductivo que ejerce la cadena alquílica que se encuentra más cercana al carbono 5. Es importante notar

que las señales mencionadas anteriormente se observan a desplazamiento químico muy semejante en todos los compuestos.

Tabla 24. Desplazamiento químico para los carbonos alifáticos de los compuestos finales

No.	NCH ₃	SCH ₂ CH ₂ N	SCH ₂ CH ₂ N	COOCH ₃	COOCH ₃	OCH ₂ CH ₃	OCH ₂ CH ₃
S1J01		32.10	45.48				
S1J02		32.03	45.40				
S1J03		32.05	45.34				
S1J04	29.87	32.55	45.27				
S1J05	30.05	32.56	45.17				
S1J06	30.08	32.53	45.17				
S1J07	30.27	32.56	45.11				
S2J08		32.04	45.38	52.00	166.75		
S2J09		31.99	45.29	52.33	165.81		
S2J10	30.11	32.51	45.21	52.00	166.63		
S2J11	30.33	32.56	45.13	52.39	165.89		
S2J12	30.16	32.54	45.14	51.97	166.71		
S2J13	30.31	32.58	45.08	52.34	165.89		
S3J14		32.64	45.90			15.21	63.86
S3J15		32.14	45.39			14.62	64.74
S3J16	29.89	32.74	45.30			14.77	63.63
S3J17	30.58	33.17	45.69			15.10	65.43
S3J18	29.9	32.58	45.3			14.73	63.42
S3J19	30.08	32.58	45.25			14.61	64.69

Las mayores variaciones en desplazamientos, como es de esperarse, se encuentran en los carbonos situados sobre el núcleo de bencimidazol (Tabla 25). El carbono de la posición 2 del bencimidazol fue el más desplazado hacia campo bajo en todos los compuestos en un rango de 149.22–156.10. Los carbonos que corresponden a la fusión de anillos (3* y 7*) son los siguientes en aparecer, en un rango de 130.26–146.28 ppm, debido a que ambos están unidos directamente a un nitrógeno cuya electronegatividad ejerce un efecto de despantallamiento. Los siguientes carbonos con mayor desplazamiento químico son los de las posiciones 5 y 6 que aparecen en un rango de 109.68–150.29 ppm, respectivamente. Estos carbonos se pueden asignar

con facilidad en la mayoría de los compuestos metilados debido al cambio significativo en desplazamiento químico comparando los compuestos sustituidos y no sustituidos en estas posiciones. Los siguientes carbonos en aparecer son los carbonos de las posiciones 4 y 7 (94.32–120.13 ppm) en los cuales el efecto del sustituyente es menos pronunciado y tienen el menor desplazamiento químico que corresponde a los carbonos aromáticos.

Tabla 25. Desplazamiento químico para los carbonos aromáticos de los compuestos finales

Clave	2'	4'	5'	2	3* y 7*	4 y 7	5 y 6
S1J01	137.55	128.5	119.42	149.22	*	111.01, 117.18	121.47**
S1J02	137.59	128.45	119.46	151.12	134.29, 136.52, 142.36, 144.25	110.21, 111.73, 116.73	121.56, 125.87
S1J03	137.59	128.47	119.42	152.87	139.2**	115.18**	123.66**
S1J04	137.53	128.53	119.37	150.65	136.79, 142.83	109.47, 117.6	121.58, 121.47
S1J05	137.51	128.51	119.32	152.17	141.55	109.66, 118.64	121.62 126.15
S1J06	137.51	128.52	119.34	152.68	135.63, 143.57	110.73, 117.05	126.03 121.49
S1J07	137.51	128.51	119.34	154.05	136.32, 142.21	111.25, 118.53	124.00, 124.02
S2J08	137.57	128.52	119.42	152.79	*	*	122.78, 122.85
S2J09	137.48	128.33	119.41	154.15	*	*	122.61, 125.00
S2J10	137.54	128.55	119.36	154.78	136.55, 146.28	111.11, 117.28	122.65, 122.79
S2J11	137.57	128.54	119.41	156.1	135.25, 145.33	112.51, 119.01	125.12 122.56
S2J12	137.52	128.52	119.36	153.55	140.11, 142.36	109.53, 118.93	122.77, 122.99
S2J13	137.54	128.53	119.38	154.7	139.30, 140.95	111.8, 120.13	122.88 125.24
S3J14	137.93	128.88	119.81	154.99	130.26, 136.42, 138.60, 144.88	95.27, 101.95, 111.72, 118.31	111.07, 147.76, 149.21
S3J15	137.47	128.40	119.36	149.53	134.93, 136.41	95.75, 110.97, 118.02	116.27
S3J16	137.54	128.51	119.41	154.81	137.23, 137.39	94.32, 118.11	110.91 148.75
S3J17	137.33	129	119.83	150.78	136.64, 137.99	95.72, 118.68	116.95 150.29
S3J18	137.47	128.49	119.32	154.53	131.34, 143.55	101.62, 111.2	150.22 109.68
S3J19	137.43	128.48	119.31	151.33	131.15, 142.19	102.49, 110.48	149.6 116.67**

*Se esperaran señales pero no se logran distinguir, no son claras o se encuentran traslapadas.

**Se observa sólo una señal que probablemente abarque dos carbonos o la segunda señal no es clara.

Es importante mencionar que algunos compuestos con hidrógeno en posición 1 y sustitución en posición 5 ó 6 presentan espectros interesantes, debido a que es posible observar las señales para dos formas tautoméricas en ^{13}C . Un ejemplo es el compuesto **S3J14** que muestra señales dobles en la región aromática para los carbonos ubicados en el anillo bencenoide. Estas señales, al compararse con los análogos metilados en posición 1 (**S3J16** y **S3J18**), sugieren que se trata de dos tautómeros (Figura 43). Por otro lado, el efecto de tautomería no es fácil de apreciar en todos los compuestos aun cuando el tiempo de adquisición del espectro es prolongado y en algunos casos se dificulta diferenciar las señales de los carbonos sobre el ruido debido a su baja intensidad.

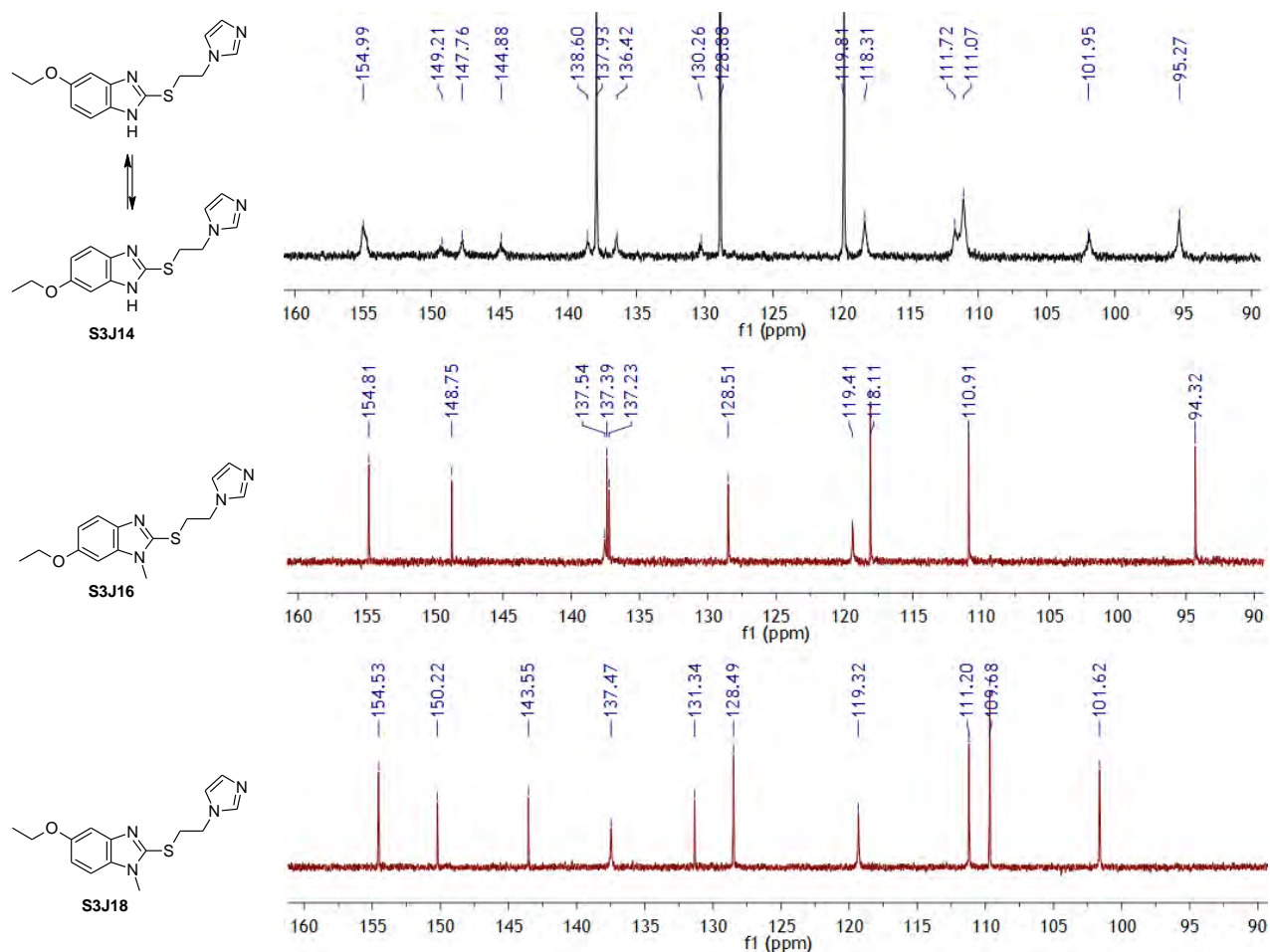


Figura 43. Efecto de la tautomería en los espectros de ^{13}C .

7.4.6.3. Espectrometría de masas y análisis elemental de los compuestos finales

El peso molecular de las moléculas sintetizadas fue determinado usando espectrometría de masas. Todas las moléculas sintetizadas se analizaron usando la técnica de impacto electrónico. Se observó el ion molecular correspondiente a cada compuesto; no obstante, es importante mencionar que los compuestos metilados en posición 1 del núcleo de bencimidazol muestran una abundancia relativa baja en comparación con sus análogos con hidrógeno. Todos los compuestos sintetizados muestran patrones de fragmentación similares. Los fragmentos más importantes coinciden con la masa para el catión radical del etilimidazol (95 u.m.a.) y para el correspondiente 2-mercaptobencimidazol (masa variable según su patrón de sustitución). En la mayor parte de las moléculas, con excepción de **S3J14** y **S3J15**, el pico base corresponde a alguno de los fragmentos mencionados. Por otra parte, para los compuestos **S3J14** y **S3J15**, el pico base corresponde a 191 y 225 u.m.a.; lo cual sugiere la pérdida de imidazol y la pérdida de la cadena alquílica en posición 5(6). Las fragmentaciones mencionadas y algunas propuestas de fragmentos adicionales cuya masa se observa en los espectros se muestran en la Figura 44.

Para todos los compuestos finales se determinó la composición elemental usando espectrometría de masas de alta resolución; los resultados se muestran en la Tabla 26. Todos los compuestos muestran una masa para el ion molecular muy cercana a la masa monoisotópica calculada para la composición esperada. Este hecho establece que los compuestos finales poseen la composición deseada.

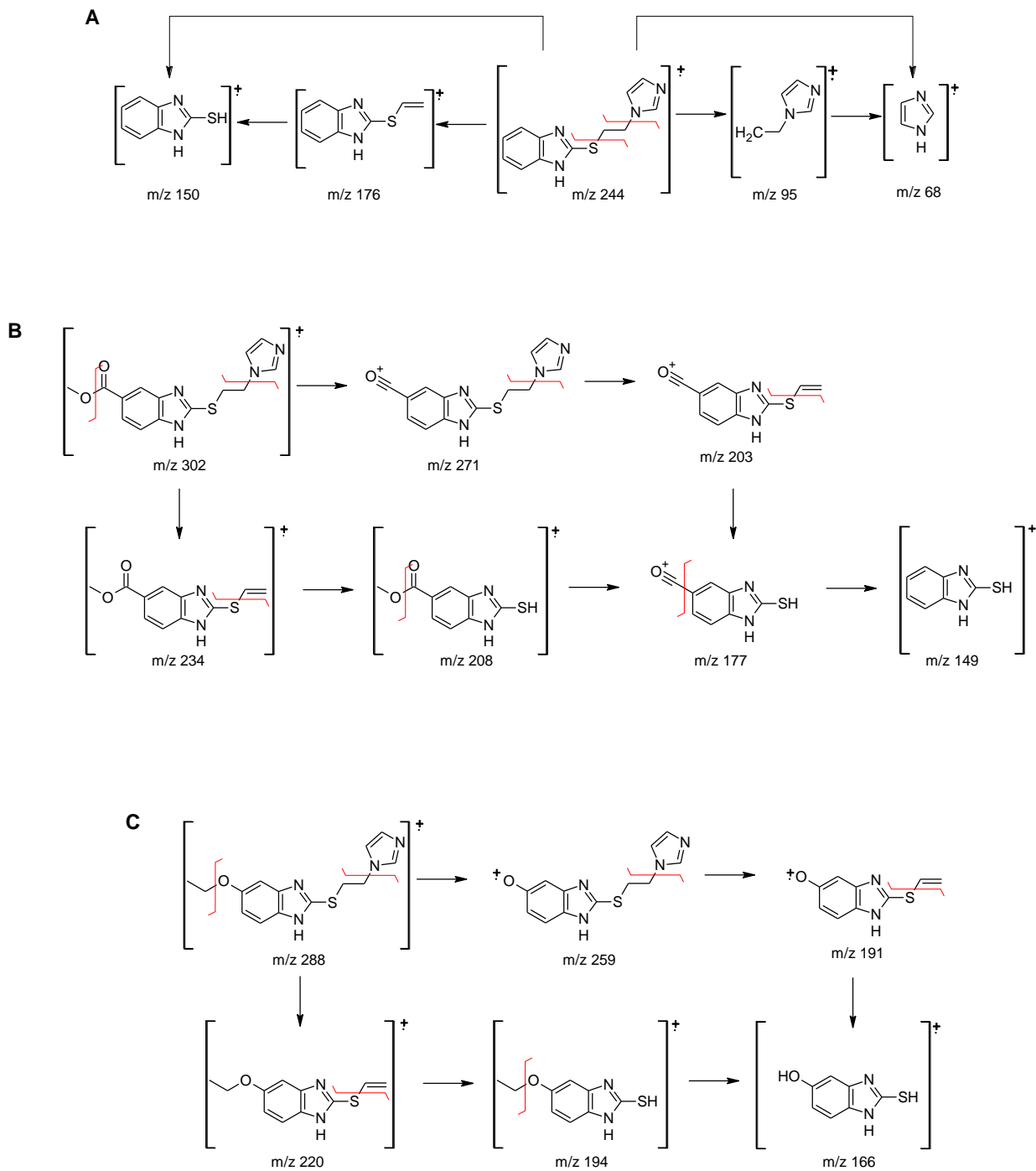


Figura 44. Fragmentos propuestos, para algunos compuestos finales representativos, mediante la interpretación de sus espectros de masas. A) Compuesto **S1J01**; B) Compuesto **S2J08**; C) Compuesto **S3J14**.

Tabla 26. Análisis elemental por espectrometría de masas de alta resolución

Clave	Composición esperada	Masa monoisotópica calculada	Masa encontrada para el ion molecular
S1J01	C ₁₂ H ₁₂ N ₄ S	244.0777	244.0791
S1J02	C ₁₂ H ₁₁ ClN ₄ S	278.0387	278.0406
S1J03	C ₁₂ H ₁₀ Cl ₂ N ₄ S	311.9998	311.9996
S1J04	C ₁₃ H ₁₄ N ₄ S	258.0934	258.0933
S1J05	C ₁₃ H ₁₃ ClN ₄ S	292.0544	292.0562
S1J06	C ₁₃ H ₁₃ ClN ₄ S	292.0544	292.0560
S1J07	C ₁₃ H ₁₂ Cl ₂ N ₄ S	326.0154	326.0172
S2J08	C ₁₄ H ₁₄ N ₄ O ₂ S	302.0832	302.0834
S2J09	C ₁₄ H ₁₃ ClN ₄ O ₂ S	336.0442	336.0422
S2J10	C ₁₅ H ₁₆ N ₄ O ₂ S	316.0988	316.0971
S2J11	C ₁₅ H ₁₅ ClN ₄ O ₂ S	350.0599	350.0604
S2J12	C ₁₅ H ₁₆ N ₄ O ₂ S	316.0988	316.0980
S2J13	C ₁₅ H ₁₅ ClN ₄ O ₂ S	350.0599	350.0622
S3J14	C ₁₄ H ₁₆ N ₄ OS	288.1039	288.1030
S3J15	C ₁₄ H ₁₅ ClN ₄ OS	322.0650	322.0631
S3J16	C ₁₅ H ₁₈ N ₄ OS	302.1196	302.1190
S3J17	C ₁₅ H ₁₇ ClN ₄ OS	336.0806	336.0800
S3J18	C ₁₅ H ₁₈ N ₄ OS	302.1196	302.1180
S3J19	C ₁₅ H ₁₇ ClN ₄ OS	336.0806	336.0819

7.4.7. Evaluación de la actividad antiprotozoaria de los compuestos finales

Los compuestos **S1J01–S3J19** se sometieron a ensayos in vitro contra los protozoarios *T. vaginalis*, *G. intestinalis* y *E. histolytica*; los resultados de actividad antiprotozoaria se muestran en la Tabla 27. Todos los compuestos evaluados mostraron mejor actividad que el metronidazol (fármaco de elección) contra los tres protozoarios. Cabe mencionar que los compuestos sintetizados presentan diferencias de potencia discretas. Sin embargo, las variaciones en actividad son suficientes para establecer algunas conclusiones sobre las relaciones estructura-actividad.

Para los compuestos **S1J01–S1J07** se aprecia un efecto favorable para la actividad antiprotozoaria al introducir cloro en posiciones 5 y 6; por ejemplo, los compuestos diclorados **S1J03** y **S1J07** tienen mejor actividad que sus análogos no clorados **S1J01** y **S1J04**, respectivamente. Adicionalmente, los compuestos **S1J01** y **S1J04** muestran que el grupo metilo favorece la actividad contra los tres protozoarios evaluados. Esta misma tendencia se observa para los compuestos **S1J03** y **S1J07** contra *T. vaginalis* pero no hay cambios significativos contra *G. intestinalis* y *E.*

histolytica. De manera general, los compuestos metilados en posición 1 muestran comparable o mejor actividad que sus análogos con hidrógeno en posición 1. Una excepción es el compuesto **S2J13** que tiene menor actividad que su análogo con hidrógeno **S2J09** contra *T. vaginalis*.

En particular, los compuestos **S3J16–S3J19**, con sustituyente etoxi, tienen los mejores valores de actividad. Es importante hacer énfasis en que el compuesto **S3J18**, el cual tiene una potente actividad giardicida, fue el único compuesto donde la actividad aumentó más de 14 veces contra *G. intestinalis* comparado con su análogo no sustituido **S1J01**.

Tabla 27. Actividad antiprotozoaria de los compuestos sintetizados

Compuesto	<i>T. vaginalis</i> Cl ₅₀ μM	<i>G. intestinalis</i> Cl ₅₀ μM	<i>E. histolytica</i> Cl ₅₀ μM
S1J01	0.1780 ± 0.0087	0.1167 ± 0.0029	0.1474 ± 0.0058
S1J02	0.1704 ± 0.0076	0.0807 ± 0.0025	0.0735 ± 0.0076
S1J03	0.1405 ± 0.0045	0.0766 ± 0.0090	0.0591 ± 0.0023
S1J04	0.1490 ± 0.0027	0.0871 ± 0.0082	0.0774 ± 0.0055
S1J05	0.1281 ± 0.0072	0.0626 ± 0.0039	0.0672 ± 0.0110
S1J06	0.1349 ± 0.0024	0.0666 ± 0.0024	0.0660 ± 0.0129
S1J07	0.0963 ± 0.0022	0.0688 ± 0.0065	0.0642 ± 0.0043
S2J08	0.1108 ± 0.0023	0.0777 ± 0.0023	0.0612 ± 0.0023
S2J09	0.0727 ± 0.0105	0.0549 ± 0.0021	0.0579 ± 0.0021
S2J10	0.1138 ± 0.0055	0.0569 ± 0.0045	0.0522 ± 0.0067
S2J11	0.0869 ± 0.0101	0.0556 ± 0.0020	0.0428 ± 0.0040
S2J12	0.0774 ± 0.0067	0.0616 ± 0.0067	0.0427 ± 0.0112
S2J13	0.0926 ± 0.0020	0.0442 ± 0.0020	0.0513 ± 0.0040
S3J14	0.0717 ± 0.0053	0.0566 ± 0.0020	0.0428 ± 0.0044
S3J15	0.0991 ± 0.0044	0.0465 ± 0.0044	0.0620 ± 0.0044
S3J16	0.0717 ± 0.0038	0.0452 ± 0.0083	0.0408 ± 0.0019
S3J17	0.0698 ± 0.0063	0.0356 ± 0.0084	0.0460 ± 0.0063
S3J18	0.0761 ± 0.0094	0.0083 ± 0.0023	0.0298 ± 0.0047
S3J19	0.0980 ± 0.0042	0.0208 ± 0.0042	0.0148 ± 0.0042
Metronidazol	0.2360 ± 0.0160	1.2260 ± 0.1250	0.3798 ± 0.1461
Albendazol	1.5905 ± 0.0113	0.0370 ± 0.0030	56.5334 ± 18.8445

La actividad tricomonocida de los nuevos compuestos se comparó con los resultados de actividad calculada con los modelos de CoMFA y CoMSIA. En la Tabla 28 se muestra la actividad experimental y la predicha con los cuatro mejores modelos expresada en valores de pCl₅₀. Las predicciones son aceptables en la mayoría de los casos; particularmente para el modelo 6, el 79% de los compuestos tienen residuales dentro de un rango de dos desviaciones estándar (2S = 0.392) lo cual se considera aceptable. Sin embargo, es difícil comparar cuál de los modelos presenta las mejores

predicciones basado en cuantos compuestos caen dentro del rango 2S debido a que el valor de S (error estándar estimado) es característico para cada modelo. Otra medida útil para concluir cuál de los modelos tuvo la mejor capacidad predictiva es el error estándar de las predicciones ($PRESS = \sum [y_{\text{experimental}} - y_{\text{calculada}}]^2$). Se calculó PRESS para los compuestos sintetizados y se observó el siguiente orden en los valores: modelo 6 (1.829) < modelo 13 (2.161) < modelo 17 (2.266) < modelo 12 (4.303). El valor de PRESS permite concluir que los modelos de CoMFA 6 y 13 tuvieron mejores aproximaciones que los modelos de CoMSIA 12 y 17. Este resultado coincide con el error estándar estimado S para los modelos de CoMFA. No obstante, aunque en el modelo de CoMSIA 17 se obtiene un valor de S más elevado las predicciones son aún mejores que en el modelo de CoMSIA 12.

Tabla 28. Predicciones para las moléculas sintetizadas usando los mejores modelos QSAR-3D (6, 12, 13 y 17) y su comparación con los resultados experimentales

	Modelo 6			Modelo 12		Modelo 13		Modelo 17	
	S = 0.196			S = 0.286		S = 0.198		S = 0.332	
	2S = 0.392			2S = 0.572		2S = 0.396		2S = 0.664	
	Exp. ^a	Calc. ^b	Res. ^c	Calc.	Res.	Calc.	Res.	Calc.	Res.
S1J01	6.75 ^d	6.82	-0.07	7.04	-0.29	6.59	0.16	7.07	-0.32
S1J02	6.77	6.95	-0.18	7.00	-0.23	6.74	0.03	7.05	-0.28
S1J03	6.85	7.31	-0.46	7.25	-0.40	7.21	-0.36	7.28	-0.43
S1J04	6.83	6.49	0.34	6.53	0.30	6.29	0.54	6.74	0.09
S1J05	6.89	6.89	0.00	6.79	0.10	6.72	0.17	6.98	-0.09
S1J06	6.87	6.55	0.32	6.50	0.37	6.41	0.46	6.72	0.15
S1J07	7.02	6.93	0.09	6.76	0.26	6.84	0.18	6.95	0.07
S2J08	6.96	6.93	0.03	6.80	0.16	6.88	0.08	6.93	0.03
S2J09	7.14	7.28	-0.14	6.90	0.24	7.16	-0.02	6.98	0.16
S2J10	6.94	6.56	0.38	6.29	0.65	6.58	0.36	6.59	0.35
S2J11	7.06	7.27	-0.21	6.57	0.49	7.24	-0.18	6.69	0.37
S2J12	7.11	6.40	0.71	6.00	1.11	6.30	0.81	6.21	0.90
S2J13	7.03	6.93	0.10	6.33	0.70	6.81	0.22	6.53	0.50
S3J14	7.14	7.16	-0.02	7.05	0.09	7.05	0.09	7.08	0.06
S3J15	7.00	7.25	-0.25	7.13	-0.13	7.21	-0.21	7.14	-0.14
S3J16	7.14	6.74	0.40	6.54	0.60	6.74	0.40	6.75	0.39
S3J17	7.16	6.85	0.31	6.64	0.52	6.87	0.29	6.81	0.35
S3J18	7.12	6.59	0.53	6.43	0.69	6.60	0.52	6.67	0.45
S3J19	7.01	6.94	0.07	6.69	0.32	7.01	0.00	6.91	0.10
PRESS ^e			1.829		4.303		2.161		2.266

^aActividad experimental.

^bActividad calculada.

^cResidual.

^dLa actividad está reportada como el $-\log CI_{50}$ donde la CI_{50} se encuentra en unidades molares.

^eError estándar de las predicciones calculado con los residuales de los compuestos sintetizados.

8. CONCLUSIONES

Los estudios de quimioinformática proporcionaron información valiosa para entender las relaciones estructura-actividad de derivados de bencimidazol con actividad tricomonocida y giardicida. Entre los resultados más importantes se encuentra la detección de *activity cliffs consenso* para *T. vaginalis* y *G. intestinalis*. Además se detectaron *activity cliffs* comunes para ambos protozoarios. Esto llevó a la conclusión de que existe un SAR heterogéneo para ambos protozoarios.

La cuantificación de *activity cliffs* (*profundos y superficiales*) para ambos protozoarios explica la pobre correlación en el modelo de CoMFA previamente reportado para *G. intestinalis* y la mejor respuesta obtenida en los de CoMFA y CoMSIA para *T. vaginalis*.

Por otra parte se encontró una relación entre la actividad tricomonocida y giardicida para varios derivados de bencimidazol, lo cual sugiere un mecanismo de acción similar contra ambos protozoarios.

Los mapas DAD, propuestos por primera vez en este trabajo, permitieron estudiar las relaciones estructura-actividad para ambos protozoarios encontrando características importantes para que las moléculas sean potentes contra ambos parásitos.

En lo que respecta a los modelos cuantitativos obtenidos mediante las metodologías de CoMFA y CoMSIA, usando bencimidazoles tricomonocidas, se observó que varios de los modelos obtenidos poseen un alto poder predictivo. Adicionalmente, los mapas de contornos obtenidos a partir de los modelos de CoMFA y CoMSIA mostraron características importantes para optimizar la actividad tricomonocida de nuevas moléculas. En general se observó que las sustituciones en posición 2 y 6 son importantes para la actividad.

A partir de las conclusiones generadas mediante los estudios computacionales, se encontró actividad antiprotozoaria para cuatro inhibidores de la bomba de protones. Particularmente el pantoprazol resultó ser un antiprotozoario potente. Estos compuestos, que actualmente son fármacos usados en enfermedades relacionadas

con la producción de ácido estomacal, pueden tener un nuevo uso como antiprotozoarios después de las pruebas clínicas correspondientes.

En lo que respecta a la síntesis de derivados de bencimidazol se obtuvieron 19 nuevos compuestos en cantidades suficientes para su caracterización y evaluación antiprotozoaria. Todos los compuestos evaluados fueron mejores antiprotozoarios que el fármaco de elección (metronidazol). Cabe destacar que los compuestos **S3J17–S3J19** están entre los compuestos más activos contra los tres protozoarios en esta serie y se encuentran entre los derivados de bencimidazol más potentes obtenidos hasta la fecha.

Al compararse la actividad tricomonocida experimental con la actividad calculada usando los modelos QSAR-3D, se observaron predicciones aceptables para la actividad tricomonocida en la mayoría de los compuestos, lo que lleva a concluir que los modelos poseen poder predictivo.

Aunque en este trabajo sólo se sintetizó un grupo representativo de 19 compuestos, los estudios de relaciones estructura-actividad presentados marcan la pauta para diseñar y sintetizar una mayor cantidad de compuestos y avanzar en la búsqueda de nuevos agentes antiprotozoarios.

9. PROCEDIMIENTOS EXPERIMENTALES

9.1 Instrumentación

El peso de las sustancias se determinó en una balanza granataria OHAUS E4000, o bien, en una balanza analítica Sartorius A210P.

Para efectuar las reacciones a temperatura constante se empleó una parrilla automática IKA modelo IKAMAG RET básica con sensor de temperatura modelo IKATRON ETS-D4 fuzzy.

La concentración de las soluciones se realizó a presión reducida por medio de un rotaevaporador Büchi R-114 con baño B-480 acoplado a una bomba de vacío marca Gast 0523-V47-6528DX ajustada a 50 mmHg y un condensador de alta eficiencia conectado a un enfriador VWR Scientific modelo 1107, o bien, acoplado a un condensador de dedo frío enfriado por un elemento refrigerante en espiral Brinkmann modelo IC30. Para disminuir la presión del sistema se utilizó una bomba de vacío marca Vacuubrand modelo PC 610 con regulador de vacío integrado CVC2.

La hidrogenación catalítica se realizó empleando un hidrogenador marca Parr Modelo 3916 EG a 60 lb/pulg² y 250 cc como máxima capacidad, utilizando un tanque de hidrógeno de la casa INFRA y como catalizador Pd/C (al 10 %) o Raney-Níquel suspendido en agua de la casa Aldrich.

El punto de fusión (pf) de los intermediarios y productos finales se determinó en un aparato Büchi B-540, utilizando capilares de vidrio y no están corregidos.

Los espectros de Resonancia Magnética Nuclear se determinaron en espectrómetros Varian modelo Unity Inova de 300 y 400 MHz. Los disolventes deuterados empleados fueron cloroformo, dimetilsulfóxido y agua deuterada para el intercambio de hidrógeno. Los desplazamientos químicos (δ) se dan en ppm. Las constantes de acoplamiento J se dan en Hertz. Los símbolos correspondientes a la multiplicidad de las señales son: s = señal simple, d = señal doble, dd = señal doble de doble, t = señal triple, m = señal múltiple, bs = señal ancha.

Los espectros de masas (EM) se realizaron por introducción directa en un espectrómetro de masas JEOL modelo JMS-SX102A empleando la técnica de impacto electrónico (IE) o mediante la técnica FAB⁺. La simbología empleada es: m/z (masa/carga), M⁺ (ión molecular), M⁺+1, M⁺+2 (picos isotópicos).

9.2. Sistemas cromatográficos

Para el desarrollo experimental de este trabajo se usaron los sistemas cromatográficos que se muestran en la Tabla 29.

Tabla 29. Sistemas cromatográficos usados en este trabajo

Sistema	Componentes	Proporción
I	Hexano-cloroformo-acetato de etilo	50:35:15
II	Hexano-acetato de etilo	60:40
III	Hexano-acetato de etilo	70:30
IV	Hexano-acetato de etilo	90:10
V	Cloroformo-metanol	80:20
VI	Cloroformo-metanol	90:10
VII	Cloroformo-metanol	95:5
VIII	Cloroformo-metanol	98:2
IX	Cloroformo-acetona	95:5
X	Acetato de etilo-metanol	80:20
XI	Hexano-acetato de etilo	50:50
XII	Diclorometano	100
XIII	Hexano-diclorometano	90:10
XIV	Hexano-diclorometano	60:40
XV	Hexano-acetato de etilo-ácido acético	60:35:5
XVI	Hexano-acetato de etilo	40:60

9.3. Extracción de los inhibidores de la bomba de protones

Para cada compuesto se molió en un mortero el contenido de las cápsulas equivalente a 200 mg de principio activo. Posteriormente, el sólido se suspendió en solución acuosa de sulfito de sodio al 1 % (20 mL) y se extrajo con acetato de etilo (20 mL). La fase orgánica se secó con sulfato de sodio anhidro, se filtró usando papel Whatman para eliminar los residuos sólidos y se evaporó el disolvente. El residuo de evaporación para el omeprazol se recristalizó usando acetonitrilo. Para el lansoprazol

y pantoprazol, los residuos de evaporación se trataron con hexano frío y éter diisopropílico, respectivamente, para generar un precipitado en ambos casos. El residuo de evaporación del rabeprazol se convirtió a su sal de sodio añadiendo un equivalente de NaOH en solución acuosa 0.1 N, se agitó hasta obtener una solución y se evaporó el agua al vacío mediante destilación azeotrópica con etanol. Finalmente, el residuo de evaporación se trató con éter diisopropílico para dar un precipitado (sal de sodio del rabeprazol).

9.3.1. *5(6)-Metoxi-2-[[[4-metoxi-3,5-dimetilpiridin-2-il]metil]sulfinil]-1H-bencimidazol (omeprazol)*

Recristalizó de acetonitrilo dando cristales blancos; pf 155–156.6 °C (descomposición). RMN ¹H (DMSO-*d*₆) δ: 2.19 (s, 3H), 2.22 (s, 3H), 3.71 (s, 3H), 3.83 (s, 3H), 4.68–4.82 (sistema AB, 2d, *J* = 13.6 Hz, 2H), 6.95 (dd, *J* = 8.9 y 2.4 Hz, 1H), 7.12 (s, 1H), 7.57 (d, *J* = 8.9 Hz, 1H), 8.21 (s, 1H), 13.47 (s, 1H); EM (FAB) *m/z*: 346 [M + 1]⁺.

9.3.2. *2-[[[3-Metil-4-(2,2,2-trifluoroetoxi)piridin-2-il]metil]sulfinil]-1H-bencimidazol (lansoprazol)*

Se obtuvo un precipitado blanco usando hexano; pf 165–166.5 °C (dec.). RMN ¹H (DMSO-*d*₆) δ: 2.20 (s, 3H), 4.76–4.89 (sistema AB, *J* = 13.8 Hz, 2H), 4.94 (q, *J* = 8.7 Hz, 1H), 7.12 (d, *J* = 5.7 Hz, 1H), 7.33 (s, 2H), 7.62 (s, 1H), 7.73 (s, 1H), 8.32 (d, *J* = 5.7 Hz, 1H), 13.62 (s, 1H); EM (FAB) *m/z*: 370 [M + 1]⁺.

9.3.3. *2-[[[4-(3-Metoxipropoxi)-3-metilpiridin-2-il]metil]sulfinil]bencimidazol, sal de sodio (rabeprazol sódico)*

Se obtuvo un precipitado blanco usando éter diisopropílico; pf 133–135 °C (descomposición). RMN ¹H (DMSO-*d*₆) δ: 1.99 (m, *J* = 6.2 Hz, 2H), 2.17 (s, 3H), 3.26 (s, 3H), 3.50 (t, *J* = 6.2 Hz, 2H), 4.10 (t, *J* = 6.2 Hz, 2H), 4.45–4.77 (sistema AB, 2d, *J* = 13.0 Hz, 2H), 6.94–6.97 (m, 3H), 7.49–7.52 (m, 2H), 8.28 (d, *J* = 5.7 Hz, 1H); EM (FAB) *m/z*: 382 [M + 1]⁺.

9.3.4. 5(6)-(Difluorometoxi)-2-[[[(3,4-dimetoxipiridin-2-il)metil]sulfinil]-1H-bencimidazol
(pantoprazol)

Se obtuvo un precipitado blanco usando éter diisopropílico; pf 137–139 °C (descomposición). RMN ¹H (DMSO-*d*₆) δ: 3.80 (s, 3H), 3.92 (s, 3H), 4.67–4.77 (sistema AB, 2d, *J* = 13.0 Hz, 2H), 7.14 (d, *J* = 5.6 Hz, 1H), 7.28 (t, *J* = 74 Hz, 1H), 7.19 (d, *J* = 8.6 Hz, 1H), 7.47 (s, 1H), 7.71 (bs, 1H), 8.18 (d, *J* = 5.5 Hz, 1H), 13.80 (s, 1H); EM (FAB) *m/z*: 384 [M + 1]⁺.

9.4. Procedimientos para la obtención de compuestos finales

9.4.1. Procedimiento general de alquilación con 1-(2-Cloroetil)-1H-imidazol (**S1J01–S1J03**, **S1J05**, **S1J06**, y **S2J08–S3J19**)

A una solución del 2-mercaptobencimidazol adecuadamente sustituido **18–34** (6 mmol) en acetonitrilo (20 mL) se agregó KOH al 50% (6.6 mmol, 1.1 eq) o DBU (6.6 mmol, 1.1 eq; para los derivados de tipo éster **23–28**). La mezcla de reacción se agitó durante 10 minutos; posteriormente se agregó 1-(2-cloroetil)-1H-imidazol (10.2 mmol, 1.7 eq) y se calentó a 60°C por un periodo de 6–8 horas. Una vez agotada la materia prima, la mezcla de reacción se llevó a pH neutro usando ácido acético y se evaporó el disolvente en el rotavapor; posteriormente, el residuo se disolvió en acetato de etilo (20–30 mL) y se lavó con agua (3 x 20 mL). La fase orgánica se secó con sulfato de sodio anhidro y se concentró en el rotavapor. El producto obtenido se purificó usando cromatografía en columna empleando acetato de etilo como fase móvil.

9.4.2. Procedimiento general de alquilación con yoduro de metilo (**S1J04** y **S1J07**)

A una solución de **S1J01** o **S1J03** (2 mmol) en acetona (5 mL) se agregó KOH al 50% (2.2 mmol, 1.1 eq) y se agitó durante 10 minutos; posteriormente se agregó yoduro de metilo (2.2 mmol, 1.1 eq) y la mezcla se agitó durante 4 horas. Una vez agotada la materia prima, se evaporó el disolvente en el rotavapor. El residuo se disolvió en acetato de etilo (20–30 mL) y se lavó con agua (3 x 20 mL). La fase orgánica se secó con sulfato de sodio anhidro y se concentró en el rotavapor hasta un

volumen de 5 mL. Finalmente, la solución obtenida se filtró sobre gel de sílice y se concentró en el rotavapor.

9.4.1.1. 2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (**S1J01**)

Se obtuvo 55% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.47 (Sistema X) y pf de 139.0–140.1 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.664 (t, 2H, $-\text{SCH}_2-$, $J = 6.45$ Hz), 4.371 (t, 2H, NCH_2- , $J = 6.45$ Hz), 6.910 (s, 1H, H-5'), 7.129 (m, 2H, H5 y H6), 7.242 (s, 1H, H-4'), 7.462 (bs, 2H; H4 y H7), 7.654 (s, 1H, H-2'), 12.654 (bs, 1H, NH); RMN ^{13}C (75.5 MHz, DMSO- d_6) δ : 32.10 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.48 ($\text{SCH}_2\text{CH}_2\text{N}$), 111.01, 117.18 (C-4 y C-7), 119.42 (C-5'), 121.47 (C-5 y C-6), 128.5 (C-4'), 137.55 (C-2'), 149.22 (C-2); EM (IE), m/z (% intensidad relativa): 244 (M^+ , 27), 176 (25), 175 (41), 150 (100), 95 (79), 94 (6), 68 (10); EMAR encontrada 244.0791 (calculada para $\text{C}_{12}\text{H}_{12}\text{N}_4\text{S}$ 244.0777).

9.4.1.2. 5-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (**S1J02**)

Se obtuvo 87% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.33 (Sistema VII) y pf de 110.8–112.5 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.667 (t, 2H, $-\text{SCH}_2-$, $J = 6.3$ Hz), 4.370 (t, 2H, NCH_2- , $J = 6.3$ Hz), 6.908 (s, 1H, H-5'), 7.148 (dd, 1H, H-6, $J = 8.6$ Hz, 1.4 Hz), 7.242 (s, 1H, H-4'), 7.455 (d, 1H, H-7, $J = 7.5$ Hz), 7.527 (s, 1H, H-4), 7.662 (s, 1H, H-2'), 12.858 (bs, 1H, NH); RMN ^{13}C (75.5 MHz, DMSO- d_6) δ : 32.03 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.40 ($\text{SCH}_2\text{CH}_2\text{N}$), 110.21, 111.73, 116.73 (C-4 y C-7), 119.46 (C-5'), 121.56, 125.87 (C-5 y C-6), 128.45 (C-4'), 134.29, 136.52, 142.3, 144.25 (C-3* y C-7*), 137.59 (C-2'), 151.12 (C-2); EM (IE), m/z (% intensidad relativa): 280 ($\text{M}^+ + 2$, 9), 278 (M^+ , 25), 211 (26), 210 (42), 209 (57), 212 (15), 186 (35), 184 (100), 95 (99), 94 (20), 68 (24); EMAR encontrada 278.0406 (calculada para $\text{C}_{12}\text{H}_{11}\text{ClN}_4\text{S}$ 278.0387).

9.4.1.3. 5,6-Dicloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (**S1J03**)

Se obtuvo 77% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.53 (Sistema V) y pf de 197.8–199.9 °C. RMN ^1H (300 MHz,

DMSO- d_6) δ : 3.665 (t, 2H, $-\text{SCH}_2-$, $J = 6.6$ Hz), 4.364 (t, 2H, NCH_2- , $J = 6.6$ Hz), 6.889 (s, 1H, H-5'), 7.227 (s, 1H, H-4'), 7.642 (s, 1H, H-2'), 7.723 (s, 1H, H-4 y H-7), 12.979 (bs, 1H, NH); RMN ^{13}C (75.5 MHz, DMSO- d_6) δ : 32.05 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.34 ($\text{SCH}_2\text{CH}_2\text{N}$), 115.18 (C-4 y C-7), 119.42 (C-5'), 123.66 (C-5 y C-6), 128.47 (C-4'), 137.59 (C-2'), 139.2 (C-3* y C-7*), 152.87 (C-2); EM (IE), m/z (% intensidad relativa): 316 ($\text{M}^+ + 4$, 4), 314 ($\text{M}^+ + 2$, 23), 312 (M^+ , 36), 248 (4), 247 (8), 246 (21), 245 (30), 244 (34), 243 (38), 218 (47), 220 (32), 222 (6), 94 (46), 95 (100), 68 (52); EMAR encontrada 311.9996 (calculada para $\text{C}_{12}\text{H}_{10}\text{Cl}_2\text{N}_4\text{S}$ 311.9998).

9.4.1.4. 2- $\{[2-(1\text{H-imidazol-1-il})\text{etil}]\text{sulfanil}\}$ -1-metil-1H-bencimidazol (**S1J04**)

Se obtuvo 86% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.31 (Sistema IV) y pf de 51.5–52.7 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.653 (s, 3H, NCH_3), 3.708 (t, 2H, $-\text{SCH}_2-$, $J = 6.3$ Hz), 4.385 (t, 2H, NCH_2- , $J = 6.45$ Hz), 6.903 (s, 1H, H-5'), 7.143–7.231 (m, 2H, H-5 y H-6), 7.248 (s, 1H, H-4'), 7.469–7.499 (m, 1H, H-7), 7.564–7.593 (m, 1H, H-4), 7.660 (s, 1H, H-2'); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 29.87 (NCH_3), 32.55 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.27 ($\text{SCH}_2\text{CH}_2\text{N}$), 109.47, 117.60 (C-4 y C-7), 119.37 (C-5'), 121.47, 121.58 (C-5 y C-6), 128.53 (C-4'), 136.79, 142.83 (C-3* y C-7*), 137.53 (C-2'), 150.65 (C-2); EM (IE), m/z (% intensidad relativa): 258 (M^+ , 4), 189 (6), 191 (1), 164 (100), 95 (89), 94 (11), 68 (42); EMAR encontrada 258.0933 (calculada para $\text{C}_{13}\text{H}_{14}\text{N}_4\text{S}$ 258.0934).

9.4.1.5. 6-Cloro-2- $\{[2-(1\text{H-imidazol-1-il})\text{etil}]\text{sulfanil}\}$ -1-metil-1H-bencimidazol (**S1J05**)

Se obtuvo 81% de rendimiento de un sólido color crema que mostró una sola mancha en ccf con R_f 0.54 (Sistema VI) y pf de 119.1–120.5 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.639 (s, 3H, N-CH_3), 3.706 (t, 2H, $-\text{SCH}_2-$, $J = 6.45$ Hz), 4.381 (t, 2H, NCH_2- , $J = 6.45$ Hz), 6.896 (s, 1H, H-5'), 7.177 (dd, 1H, H-5, $J = 8.4$, $J = 2.1$ Hz), 7.240 (s, 1H, H-4'), 7.557 (d, 1H, H-4, $J = 8.4$ Hz), 7.655 y 7.661 (2s, 2H, H-7 y H-2'); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 30.05 (NCH_3), 32.56 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.17 ($\text{SCH}_2\text{CH}_2\text{N}$), 109.66, 118.64 (C-4 y C-7), 119.32 (C-5'), 121.62 (C-5), 126.15 (C-6), 128.51 (C-4'), 137.51 (C-2'), 141.55 (C-3*), 152.17 (C-2); EM (IE), m/z (% Intensidad relativa): 294

($M^+ + 2$, 2), 292 (M^+ , 6), 226 (1), 225 (4), 224 (2), 223 (6), 200 (26), 198 (75), 95 (100), 94 (13), 68 (15); EMAR encontrada 292.0562 (calculada para $C_{13}H_{13}ClN_4S$ 292.0544).

9.4.1.6. 5-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (**S1J06**)

Se obtuvo 71% de rendimiento de un sólido rosa pálido que mostró una sola mancha en ccf con R_f 0.33 (Sistema VI) y pf de 110.2–111.7 °C. RMN 1H (300 MHz, DMSO- d_6) δ : 3.653 (s, 3H, NCH_3), 3.712 (t, 2H, $-SCH_2-$, $J = 6.45$ Hz), 4.376 (t, 2H, NCH_2- , $J = 6.45$ Hz), 6.887 (s, 1H, H-5'), 7.224 (dd, 1H, H-6, $J = 8.7$ Hz, $J = 1.8$ Hz), 7.238 (s, 1H, H-4'), 7.525 (d, 1H, H-7, $J = 8.7$ Hz), 7.630 (d, 1H, H-4, $J = 1.8$ Hz), 7.649 (s, 1H, H-2'); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 30.08 (NCH_3), 32.53 (SCH_2CH_2N), 45.17 (SCH_2CH_2N), 110.73, 117.05 (C-4 y C-7), 119.34 (C-5'), 121.49 (C-6), 126.03 (C-5), 128.52 (C-4'), 135.63, 143.57 (C-3* y C-7*), 137.51 (C-2'), 152.68 (C-2); EM (IE), m/z (% Intensidad relativa): 292 (M^+ , 6), 294 ($M^+ + 2$, 2), 226 (1), 225 (5), 224 (2), 223 (8), 200 (33), 198 (93), 95 (100), 94 (14), 68 (15); EMAR encontrada 292.0560 (calculada para $C_{13}H_{13}ClN_4S$ 292.0544).

9.4.1.7. 5,6-Dicloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (**S1J07**)

Se obtuvo 84% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.37 (Sistema VI) y pf de 142.5–144.0 °C. RMN 1H (300 MHz, DMSO- d_6) δ : 3.633 (s, 3H, NCH_3), 3.709 (t, 2H, $-SCH_2-$, $J = 6.45$ Hz), 4.376 (t, 2H, NCH_2- , $J = 6.3$ Hz), 6.885 (s, 1H, H-5'), 7.234 (s, 1H, H-4'), 7.652 (s, 1H, H-2'), 7.802 (s, 1H, H-7), 7.867 (s, 1H, H-4); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 30.27 (NCH_3), 32.56 (SCH_2CH_2N), 45.11 (SCH_2CH_2N), 111.25, 118.53 (C-4 y C-7), 119.34 (C-5'), 124.00, 124.02 (C-5 y C-6), 128.51 (C-4'), 136.32, 142.21 (C-3* y C-7*), 137.51 (C-2'), 154.05 (C-2); EM (IE), m/z (% Intensidad relativa): 330 ($M^+ + 4$, 1), 328 ($M^+ + 2$, 2), 326 (M^+ , 3), 260 (1), 259 (3), 258 (1), 257 (3), 236 (5), 234 (26), 232 (39), 95 (100), 94(8), 68(12). EMAR encontrada 326.0172 (calculada para $C_{13}H_{12}Cl_2N_4S$ 326.0154).

9.4.1.8. 2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol-5-carboxilato de metilo

(S2J08)

Se obtuvo 85% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.60 (Sistema V) y pf de 182.3–183.5 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.702 (t, 2H, $-\text{SCH}_2-$, $J = 6.45$ Hz), 3.851 (s, 3H, $-\text{COOCH}_3$), 4.395 (t, 2H, NCH_2- , $J = 6.45$ Hz), 6.905 (s, 1H, H-5'), 7.246 (s, 1H, H-4'), 7.537 (d, 1H, H-7, $J = 8.4$ Hz), 7.668 (s, 1H, H-2'), 7.785 (dd, 1H, H-6, $J = 8.55$ Hz, $J = 1.65$ Hz), 8.047 (s, 1H, H-4), 12.902 (bs, 1H, NH); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 32.04 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.38 ($\text{SCH}_2\text{CH}_2\text{N}$), 52.00 (COOCH_3), 119.42 (C-5'), 122.78, 122.85 (C-5 y C-6), 128.52 (C-4'), 137.57 (C-2'), 152.79 (C-2), 166.75 (COOCH_3); EM (IE), m/z (% Intensidad relativa): 302 (M^+ , 32), 271 (3), 234 (19), 233 (36), 208 (44), 203 (18), 177 (31), 149 (11), 95 (100), 94 (62), 68 (52); EMAR encontrada 302.0834 (calculada para $\text{C}_{14}\text{H}_{14}\text{N}_4\text{O}_2\text{S}$ 302.0832).

9.4.1.9. 6-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol-5-carboxilato de

metilo (S2J09)

Se obtuvo 69% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.42 (Sistema VI) y pf de 197.5–198.9 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.690 (t, 2H, $-\text{SCH}_2-$, $J = 6.45$ Hz), 3.850 (s, 3H, $-\text{COOCH}_3$), 4.381 (t, 2H, NCH_2- , $J = 6.45$ Hz), 6.896 (s, 1H, H-5'), 7.240 (s, 1H, H-4'), 7.633 (s, 1H, H-7), 7.660 (s, 1H, H-2'), 7.909 (s, 1H, H-4), 12.908 (bs, 1H, NH); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 31.99 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.29 ($\text{SCH}_2\text{CH}_2\text{N}$), 52.33 (COOCH_3), 119.41 (C-5'), 122.61, 125.00 (C-5 y C-6), 128.33 (C-4'), 137.48 (C-2'), 154.15 (C-2), 165.81 (COOCH_3); EM (IE), m/z (% Intensidad relativa): 338 ($\text{M}^+ + 2$, 12), 336 (M^+ , 31), 305 (6), 307 (2), 270 (11), 269 (20), 268 (29), 267 (42), 270 (11), 242 (34), 244 (13), 237 (29), 239 (11), 211 (34), 213 (13), 183 (9), 95 (100), 94(7), 68 (12); EMAR encontrada 336.0422 (calculada para $\text{C}_{14}\text{H}_{13}\text{ClN}_4\text{O}_2\text{S}$ 336.0442).

9.4.1.10. 2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-6-carboxilato de metilo (**S2J10**)

Se obtuvo 60% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.42 (Sistema VI) y pf de 119.6–120.5 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.706 (s, 3H, NCH₃), 3.749 (t, 2H, –SCH₂–, J = 6.45 Hz), 3.871 (s, 3H, –COOCH₃), 4.401 (t, 2H, NCH₂–, J = 6.45 Hz), 6.895 (s, 1H, H-5'), 7.252 (s, 1H, H-4'), 7.634 (d, 1H, H-4, J = 8.4 Hz), 7.667 (s, 1H, H-2'), 7.810 (dd, 1H, H-5, J = 8.55 Hz, J = 1.65 Hz), 8.094 (d, 1H, H7, J = 1.2 Hz); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 30.11 (NCH₃), 32.51 (SCH₂CH₂N), 45.21 (SCH₂CH₂N), 52.00 (COOCH₃), 111.11, 117.28 (C-4 y C-7), 119.36 (C-5'), 122.65, 122.79 (C-5 y C-6), 128.55 (C-4'), 136.55, 146.28 (C-3* y C-7*), 137.54 (C-2'), 154.78 (C-2), 166.63 (COOCH₃); EM (IE), m/z (% Intensidad relativa): 316 (M⁺, 10), 285 (7), 248 (3), 247 (8), 222 (100), 191 (36), 163 (12), 95 (59), 94 (8), 68 (6); EMAR encontrada 316.0971 (calculada para C₁₅H₁₆N₄O₂S 316.0988).

9.4.1.11. 5-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-6-carboxilato de metilo (**S2J11**)

Se obtuvo 63% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.33 (Sistema VI) y pf de 136.8–138.8 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.685 (s, 3H, NCH₃), 3.741 (t, 2H, –SCH₂–, J = 6.45 Hz), 3.871 (s, 3H, –COOCH₃), 4.388 (t, 2H, NCH₂–, J = 6.45 Hz), 6.881 (s, 1H, H-5'), 7.243 (s, 1H, H-4'), 7.667 (s, 1H, H-2'), 7.733 (s, 1H, H-4), 8.011 (s, 1H, H-7); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 30.33 (NCH₃), 32.56 (SCH₂CH₂N), 45.13 (SCH₂CH₂N), 52.39 (COOCH₃), 112.51, 119.01 (C-4 y C-7), 119.41 (C-5'), 122.56 (C-6), 125.12 (C-5), 128.54 (C-4'), 135.25, 145.33 (C-3* y C-7*), 137.57 (C-2'), 156.10 (C-2), 165.89 (COOCH₃); EM (IE), m/z (% Intensidad relativa): 352 (M+ + 2, 2), 350 (M+, 5), 321 (3), 319 (9), 284 (1), 283 (3), 282 (2), 281 (6), 258 (32), 256 (94), 227 (12), 225 (36), 197 (11), 95 (100), 68 (3); EMAR encontrada 350.0604 (calculada para C₁₅H₁₅ClN₄O₂S 350.0599).

9.4.1.12. 2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-5-carboxilato de metilo (**S2J12**)

Se obtuvo 52% de rendimiento de un sólido rosa pálido que mostró una sola mancha en ccf con R_f 0.32 (Sistema VI) y pf de 116.3–117.9 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.688 (s, 3H, NCH₃), 3.737 (t, 2H, –SCH₂–, J = 6.45 Hz), 3.862 (s, 3H, –COOCH₃), 4.401 (t, 2H, NCH₂–, J = 6.3 Hz), 6.890 (s, 1H, H-5'), 7.256 (s, 1H, H-4'), 7.598 (d, 1H, H-7, J = 8.4 Hz), 7.670 (s, 1H, H-2'), 7.842 (dd, 1H, H-6, J = 8.55 Hz, J = 1.65 Hz), 8.143 (d, 1H, H4, J = 1.2 Hz); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 30.16 (NCH₃), 32.54 (SCH₂CH₂N), 45.14 (SCH₂CH₂N), 51.97 (COOCH₃), 109.53, 118.93 (C-4 y C-7), 119.36 (C-5'), 122.77, 122.99 (C-5 y C-6), 128.52 (C-4'), 137.52 (C-2'), 140.11, 142.36 (C-3* y C-7*), 153.55 (C-2), 166.71 (COOCH₃); EM (IE), m/z (% Intensidad relativa): 350 (M⁺, 1), 285 (2), 247 (2), 222 (29), 191 (11), 95 (100), 94 (7), 68 (10); EMAR encontrada 316.0980 (calculada para C₁₅H₁₆N₄O₂S 316.0988).

9.4.1.13. 6-Cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-5-carboxilato de metilo (**S2J13**)

Se obtuvo 70% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.56 (Sistema V) y pf de 126.9–128.3 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.662 (s, 3H, NCH₃), 3.722 (t, 2H, –SCH₂–, J = 6.3 Hz), 3.856 (s, 3H, –COOCH₃), 4.389 (t, 2H, NCH₂–, J = 6.3 Hz), 6.884 (s, 1H, H-5'), 7.245 (s, 1H, H-4'), 7.663 (s, 1H, H-2'), 7.788 (s, 1H, H-7), 7.996 (s, 1H, H-4); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 30.31 (NCH₃), 32.58 (SCH₂CH₂N), 45.08 (SCH₂CH₂N), 52.34 (COOCH₃), 111.80, 120.13 (C-4 y C-7), 119.38 (C-5'), 122.88 (C-5), 125.24 (C-6), 128.53 (C-4'), 137.54 (C-2'), 139.30, 140.95 (C-3* y C-7*), 154.70 (C-2), 165.89 (COOCH₃); EM (IE), m/z (% Intensidad relativa): 352 (M⁺ + 2, 3), 350 (M⁺, 9), 321 (4), 319 (13), 283 (5), 281 (7), 258 (35), 256 (100), 227 (15), 225 (46), 95 (93), 68 (3); EMAR encontrada 350.0622 (calculada para C₁₅H₁₅ClN₄O₂S 350.0599).

9.4.1.14. 5-Etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (**S3J14**)

Se obtuvo 80% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.53 (cloroformo/metanol 80:20) y pf de 96–105 °C. RMN ^1H (400 MHz, DMSO- d_6) δ : 1.329 (t, 3H, $\text{CH}_3\text{CH}_2\text{O}-$, $J = 7.2$ Hz), 3.615 (t, 2H, $-\text{SCH}_2-$, $J = 6.6$ Hz), 4.013 (q, 2H, $\text{CH}_3\text{CH}_2\text{O}-$, $J = 7.06$ Hz), 4.347 (t, 2H, NCH_2- , $J = 6.6$ Hz), 6.742 (dd, 1H, H-6, $J = 9.2$, Hz, $J = 2.8$ Hz), 6.903 (s, 1H, H-5'), 7.080 (bs, 1H, H-4), 7.233 (s, 1H, H-4'), 7.397 (bs, 1H, H-7), 7.649 (s, 1H, H-2'), 12.458 (s, 1H, NH); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 15.21 (OCH_2CH_3), 32.64 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.90 ($\text{SCH}_2\text{CH}_2\text{N}$), 63.86 (OCH_2CH_3), 95.27, 101.95, 111.72, 118.31 (C-4 y C-7), 111.07, 147.76, 149.21 (C-5 y C-6), 119.81 (C-5'), 128.88 (C-4'), 130.26, 136.42, 138.60, 144.88 (C-3* y C-7*), 137.93 (C-2'), 154.99 (C-2); EM (IE), m/z (% Intensidad relativa): 288 (M^+ , 32), 259 (3), 220 (73), 219 (42), 194 (47), 192 (45), 191 (100), 166 (65), 165 (56), 95 (97), 94 (12), 68 (26); EMAR encontrada 288.1030 (calculada para $\text{C}_{14}\text{H}_{16}\text{N}_4\text{OS}$ 288.1039).

9.4.1.15. 6-Cloro-5-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (**S3J15**)

Se obtuvo 82% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.40 (Sistema X) y pf de 123.4–126.0 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 1.367 (t, 3H, $\text{CH}_3\text{CH}_2\text{O}-$, $J = 7.05$ Hz), 3.627 (t, 2H, $-\text{SCH}_2-$, $J = 6.45$ Hz), 4.092 (q, 2H, $\text{CH}_3\text{CH}_2\text{O}-$, $J = 7.00$ Hz), 4.351 (t, 2H, NCH_2- , $J = 6.45$ Hz), 6.899 (s, 1H, H-5'), 7.071 (bs, 1H, H-4), 7.230 (s, 1H, H-4'), 7.757 (bs, 1H, H-7), 7.647 (s, 1H, H-2'), 12.644 (s, 1H, NH); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 14.62 (OCH_2CH_3), 32.14 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.39 ($\text{SCH}_2\text{CH}_2\text{N}$), 64.74 (OCH_2CH_3), 95.75, 110.97, 118.02 (C-4 y C-7), 116.27 (C-5 y C-6), 119.36 (C-5'), 128.40 (C-4'), 134.93, 136.41 (C-3* y C-7*), 137.47 (C-2'), 149.53 (C-2); EM (IE), m/z (% Intensidad relativa): 324 ($\text{M}^+ + 2$, 22), 322 (M^+ , 63), 295 (3), 293 (8), 256 (25), 255 (17), 254 (71), 253(24), 230 (9), 228 (42), 227 (42), 226 (54), 225 (100), 202 (17), 201 (15), 200 (51), 199 (34), 95 (56), 94 (6), 68 (13); EMAR encontrada 322.0631 (calculada para $\text{C}_{14}\text{H}_{15}\text{ClN}_4\text{OS}$ 322.0650).

9.4.1.16. 6-Etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (**S3J16**)

Se obtuvo 90% de rendimiento de un sólido amarillo claro que mostró una sola mancha en ccf con R_f 0.55 (Sistema X) y pf de 53.0–57.1 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 1.347 (t, 3H, $\text{CH}_3\text{CH}_2\text{O}$ -, $J = 7.1$ Hz), 3.612 (s, 3H, NCH_3), 3.645 (t, 2H, $-\text{SCH}_2$ -, $J = 6.3$ Hz), 4.055 (q, 2H, $\text{CH}_3\text{CH}_2\text{O}$ -, $J = 7.0$ Hz), 4.354 (t, 2H, NCH_2 -, $J = 6.3$ Hz), 6.772 (dd, 1H, H-5, $J = 8.85$, $J = 2.25$ Hz), 6.898 (s, 1H, H-5'), 7.064 (d, 1H, H-7, $J = 2.4$ Hz), 7.239 (s, 1H, H-4'), 7.439 (d, 1H, H4, $J = 8.7$ Hz), 7.655 (s, 1H, H-2'); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 14.77 (OCH_2CH_3), 29.89 (NCH_3), 32.74 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.30 ($\text{SCH}_2\text{CH}_2\text{N}$), 63.63 (OCH_2CH_3), 94.32, 110.91 (C-5), 118.11 (C-4 y C-7), 119.41 (C-5'), 128.51 (C-4'), 137.23, 137.39 (C-3* y C-7*), 137.54 (C-2'), 148.75 (C-6), 154.81 (C-2); EM (IE), m/z (% Intensidad relativa): 302 (M^+ , 29), 273 (3), 235 (3), 234 (2), 233 (4), 208 (100), 180 (42), 179 (34), 95 (21), 94 (2), 68 (5); EMAR encontrada 302.1190 (calculada para $\text{C}_{15}\text{H}_{18}\text{N}_4\text{OS}$ 302.1196).

9.4.1.17. 5-Cloro-6-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (**S3J17**)

Se obtuvo 75% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.50 (Sistema X) y pf de 124.4–126.5 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 1.390 (t, 3H, $\text{CH}_3\text{CH}_2\text{O}$ -, $J = 6.9$ Hz), 3.633 (s, 3H, NCH_3), 3.661 (t, 2H, $-\text{SCH}_2$ -, $J = 6.3$), 4.138 (q, 2H, $\text{CH}_3\text{CH}_2\text{O}$ -, $J = 7.00$ Hz), 4.356 (t, 2H, NCH_2 -, $J = 6.45$ Hz), 6.889 (s, 1H, H-5'), 7.230 (s, 1H, H-4'), 7.288 (s, 1H, H-7), 7.622 (s, 1H, H-4), 7.645 (s, 1H, H-2'); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 15.10 (OCH_2CH_3), 30.58 (NCH_3), 33.17 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.69 ($\text{SCH}_2\text{CH}_2\text{N}$), 65.43 (OCH_2CH_3), 95.72, 118.68 (C-4 y C-7), 116.95 (C-5), 119.83 (C-5'), 129.00 (C-4'), 136.64, 137.99 (C-3* y C-7*), 137.33 (C-2'), 150.29 (C-6), 150.78 (C-2); EM (IE), m/z (% Intensidad relativa): 338 ($\text{M}^+ + 2$, 13), 336 (M^+ , 35), 309 (2), 307 (6), 270 (2), 269 (3), 268 (4), 267 (4), 244 (35), 242 (100), 216 (28), 215 (20), 214 (79), 213 (28), 95 (71), 94 (4), 68 (12); EMAR encontrada 336.0800 (calculada para $\text{C}_{15}\text{H}_{17}\text{ClN}_4\text{OS}$ 336.0806).

9.4.1.18. 5-Etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (**S3J18**)

Se obtuvo 72% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.40 (Sistema X) y pf de 111.1–113.3 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 1.334 (t, 3H, $\text{CH}_3\text{CH}_2\text{O}$ -, $J = 6.9$ Hz), 3.616 (s, 3H, NCH_3), 3.674 (t, 2H, $-\text{SCH}_2$ -, $J = 6.45$ Hz), 4.026 (q, 2H, $\text{CH}_3\text{CH}_2\text{O}$ -, $J = 6.9$ Hz), 4.371 (t, 2H, NCH_2 -, $J = 6.45$ Hz), 6.815 (dd, 1H, H-6, $J = 8.7$ Hz, $J = 2.4$ Hz), 6.907 (s, 1H, H-5'), 7.123 (d, 1H, H-4, $J = 2.4$ Hz), 7.244 (s, 1H, H-4'), 7.352 (d, 1H, H-7, $J = 9.0$ Hz), 7.661 (s, 1H, H-2'); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 14.73 (OCH_2CH_3), 29.90 (NCH_3), 32.58 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.30 ($\text{SCH}_2\text{CH}_2\text{N}$), 63.42 (OCH_2CH_3), 101.62, 111.20 (C-4 y C-7), 109.68 (C-6), 119.32 (C-5'), 128.49 (C-4'), 131.34, 143.55, (C-3* y C-7*), 137.47 (C-2'), 150.22 (C-5), 154.53 (C-2); EM (IE), m/z (% Intensidad relativa): 302 (M^+ , 38), 273 (2), 235 (3), 234 (4), 233 (6), 208 (100), 180 (40), 179 (35), 95 (7), 68 (2). EMAR encontrada 302.1180 (calculada para $\text{C}_{15}\text{H}_{18}\text{N}_4\text{OS}$ 302.1196).

9.4.1.19. 6-Cloro-5-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (**S3J19**)

Se obtuvo 70% de rendimiento de un sólido blanco que mostró una sola mancha en ccf con R_f 0.30 (Sistema X) y pf de 139.7–142.2 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 1.370 (t, 3H, $\text{CH}_3\text{CH}_2\text{O}$ -, $J = 7.05$ Hz), 3.616 (s, 3H, NCH_3), 3.678 (t, 2H, $-\text{SCH}_2$ -, $J = 6.45$ Hz), 4.105 (q, 2H, $\text{CH}_3\text{CH}_2\text{O}$ -, $J = 6.9$ Hz), 4.374 (t, 2H, NCH_2 -, $J = 6.45$ Hz), 6.899 (s, 1H, H-5'), 7.240 (s, 1H, H-4'), 7.314 (s, 1H, H-4), 7.660 (s, 2H, H-7 y H-2'); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 14.61 (OCH_2CH_3), 30.08 (NCH_3), 32.58 ($\text{SCH}_2\text{CH}_2\text{N}$), 45.25 ($\text{SCH}_2\text{CH}_2\text{N}$), 64.69 (OCH_2CH_3), 102.49, 110.48 (C-4 y C-7), 116.67 (C-6), 119.31 (C-5'), 128.48 (C-4'), 131.15, 142.19 (C-3* y C-7*), 137.43 (C-2'), 149.60 (C-5), 151.33 (C-2); EM (IE), m/z (% Intensidad relativa): 338 ($\text{M}^+ + 2$, 15), 336 (M^+ , 41), 307 (2), 270 (2), 269 (4), 268 (4), 267 (5), 244 (36), 242 (100), 216 (28), 215 (21), 214 (77), 213 (37), 95 (32), 94 (2), 68 (7); EMAR encontrada 336.0819 (calculada para $\text{C}_{15}\text{H}_{17}\text{ClN}_4\text{OS}$ 336.0806).

9.5. Procedimiento de obtención de 2-mercaptobencimidazoles

9.5.1. Procedimiento general (18–34)

Una mezcla que contiene etanol (50 mL), agua (5 mL), hidróxido de potasio (0.0192 mol, 1.2 eq) y disulfuro de carbono (0.0192 mol, 1.2 eq) se agitó a temperatura ambiente durante 15 minutos. Posteriormente se agregó el 1,2-diaminobenceno (*o*-fenilendiamina) apropiadamente sustituido **1–17** (0.0160 mol) y se calentó a 50 °C bajo atmósfera de nitrógeno durante toda la noche. Una vez agotada la materia prima, la mezcla de reacción se vertió sobre agua helada (250 mL) y se neutralizó con ácido acético. El precipitado formado se separó por filtración al vacío y se recrystalizó usando carbón activado [70].

9.5.1.1. 2-Mercapto-1*H*-bencimidazol (**18**)

El producto se recrystalizó de etanol dando un 85% de rendimiento de cristales blancos que mostraron ser un solo compuesto en ccf con R_f 0.53 (Sistema VI) y pf de 301–304.6 °C (lit: 303–304) [70].

9.5.1.2. 5-Cloro-2-mercapto-1*H*-bencimidazol (**19**)

El producto se recrystalizó de etanol-agua para dar un 56% de rendimiento de un sólido amarillo pálido que mostró ser un solo compuesto en ccf con R_f 0.52 (Sistema VII) y pf de 327.0–329.0 °C. (lit: 327–329 °C) [11].

9.5.1.3. 5,6-Dicloro-2-mercapto-1*H*-bencimidazol (**20**)

El producto se recrystalizó de metanol dando un 73% de rendimiento de un sólido blanco que mostró ser un solo compuesto en ccf con R_f 0.63 (Sistema XI) y pf de 344–346.1 °C (lit: 344 °C) [11, 99].

9.5.1.4. 6-Cloro-1-metil-2-mercapto-1H-bencimidazol (**21**)

El producto se recristalizó de etanol-acetato de etilo dando un 67% de rendimiento de cristales beige que mostraron ser un solo compuesto por ccf con R_f 0.59 (Sistema VII) con pf de 280.0–282.5 °C (lit: 273–275 °C) [100].

9.5.1.5. 5-Cloro-1-metil-2-mercapto-1H-bencimidazol (**22**)

El producto se recristalizó de etanol-acetato de etilo dando un 50% de rendimiento de cristales blancos en forma de agujas que mostraron ser un solo compuesto por ccf con R_f 0.4 (Sistema I) con pf de 233.0–234.3 °C (lit: 230–231 °C) [100].

9.5.1.6. 2-Mercapto-1H-bencimidazol-5-carboxilato de metilo (**23**)

El producto se recristalizó de etanol-agua para dar un 62% de rendimiento de cristales amarillo claro que mostraron ser un solo compuesto en ccf con R_f 0.47 (Sistema IV) con pf de 306.1–309.0 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.821 (s, 3H, –COOCH₃), 7.202 (dd, 1H, H-7, $J = 8.39$, $J = 0.59$ Hz), 7.631 (dd, 1H, H-4, $J = 1.49$, $J = 0.59$ Hz), 7.745 (dd, 1H, H-6, $J = 8.39$ Hz, $J = 1.79$ Hz); 12.838 (sa, 2H, NH y SH o 2NH); RMN ^{13}C (75.5 MHz, DMSO- d_6) δ : 52.117 (–COOCH₃), 109.267 (C-7), 110.06 (C-4), 123.50 (C-5), 124.13 (C-6), 132.24 (C-3a), 135.93 (C-7a), 166.12 (COOCH₃), 170.17 (C-2); EM, m/z (% intensidad relativa): 208 (M^+ , 100), 177 (67), 149 (21).

9.5.1.7. 6-Cloro-2-mercapto-1H-bencimidazol-5-carboxilato de metilo (**24**)

El producto se recristalizó de propanol para dar un 57% de rendimiento como polvo color salmón que mostró ser un solo compuesto en ccf con R_f 0.38 (Sistema VI) con pf de 243.9–244.8 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.824 (s, 3H, –COOCH₃), 7.211 (s, 1H, H-4), 7.518 (s, 1H, H-7), 12.911 (s, 2H, NH y SH o 2NH). RMN ^{13}C (75.5 MHz, DMSO- d_6) δ : 52.44 (–COOCH₃), 111.06 y 111.55 (C-4 y C-7), 122.68 (C-5), 126.56 (C-3a), 131.08 (C-6), 135.40 (C-7a), 165.08 (–COOCH₃), 170.99 (C-2). EM, m/z (% intensidad relativa): 244 (M^{+2} , 38), 242 (M^+ , 100), 211 (82), 183 (21).

9.5.1.8. 1-Metil-2-mercapto-1H-bencimidazol-6-carboxilato de metilo (25)

El producto se recristalizó de acetato de etilo para dar un 56% de rendimiento de sólido blanco que mostró ser un solo compuesto en ccf con R_f 0.67 (Sistema VI) y con pf de 243.2–243.8 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.688 (s, 3H, NCH₃), 3.860 (s, 3H, –COOCH₃), 7.248 (d, 1H, H-7, $J = 8.1$ Hz), 7.815 (dd, 1H, H-6, $J = 8.6$ y $J = 1.5$ Hz), 7.877 (d, 1H, H-4, $J = 1.5$ Hz), 13.088 (s, 1H, SH o NH). RMN ^{13}C (100 MHz, DMSO- d_6) δ : 30.45 (NCH₃), 52.49 (–COOCH₃), 109.65 y 110.57 (C-4 y C-7), 123.79 y 124.94 (C-5 y C-6), 133.36 (C-7a), 134.52 (C-3a), 166.54 (COOCH₃), 170.49 (C-2); EM, m/z (% intensidad relativa): 222 (M^+ , 100 %), 191 (42), 163 (22).

9.5.1.9. 5-Cloro-1-metil-2-mercapto-1H-bencimidazol-6-carboxilato de metilo (26)

El producto se recristalizó de etanol para dar un 42% de rendimiento como un sólido beige que mostró ser un solo compuesto en ccf con R_f 0.64 (Sistema VII) y pf de 230.5–231.4 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 3.644 (s, 3H, NCH₃), 3.860 (s, 3H, –COOCH₃), 7.244 (s, 1H, H-4), 7.785 (s, 1H, H-7), 13.104 (s, 1H, SH o NH); RMN ^{13}C (75.5 MHz, DMSO- d_6) δ : 30.30 (NCH₃), 52.52 (COOCH₃), 111.05 y 111.78 (C-4 y C-7), 123.03 (C-6), 126.84 (C-5), 131.96 (C-7a), 133.66 (C-3a), 165.32 (COOCH₃), 171.06 (C-2); EM, m/z (% intensidad relativa): 258 (M^{+2} , 36), 256 (M^+ , 100), 227 (22), 225 (62), 199 (12), 197 (32).

9.5.1.10. 1-Metil-2-mercapto-1H-bencimidazol-5-carboxilato de metilo (27)

El producto se recristalizó de etanol para dar un 77% de rendimiento de un sólido blanco que mostró ser un solo compuesto en ccf con R_f 0.5 (Sistema VII) y con pf de 251.4–252.3 °C. ^1H RMN (300 MHz, DMSO- d_6) δ : 3.661 (s, 3H, NCH₃), 3.847 (s, 3H, –COOCH₃), 7.44 (d, 1H, H-7, $J = 8.4$ Hz), 7.663 (d, 1H, H-4, $J = 1.5$ Hz), 7.809 (dd, 1H, H-6, $J = 8.4$ Hz, $J = 1.5$ Hz), 12.993 (s, 1H, SH o NH); ^{13}C RMN (75.5 MHz, DMSO- d_6) δ : 30.28 (NCH₃), 51.20 (COOCH₃), 109.365 y 110.115 (C-4 y C-7), 123.805 y 123.941 (C-5 y C-6), 130.629 (C-7a), 136.584 (C-3a), 166.083 (COOCH₃), 170.513 (C-2); EM, m/z (% intensidad relativa): 208 (M^+ , 100), 177 (67), 149 (21).

9.5.1.11. 6-Cloro-1-metil-2-mercapto-1H-bencimidazol-5-carboxilato de metilo (**28**)

El producto se recristalizó de etanol-acetato de etilo para dar un 65% de rendimiento de un sólido blanco que mostró ser un solo compuesto en ccf con R_f 0.77 (Sistema VII) y pf de 250.2–251.8 °C. ^1H RMN (300 MHz, DMSO- d_6) δ : 3.631 (s, 3H, NCH₃), 3.841 (s, 3H, –COOCH₃), 7.535 (s, 1H, H-7), 7.619 (s, 1H, H-4), 13.096 (s, 1H, SH o NH); ^{13}C RMN (75.5 MHz, DMSO- d_6) δ : 30.42 (NCH₃), 52.49 (COOCH₃), 111.57 y 111.59 (C-4 y C-7), 123.13 (C-5), 126.57 (C-6), 129.32 (C-7a), 136.09 (C-3a), 165.04 (COOCH₃), 171.23 (C-2); EM, m/z (% intensidad relativa): 244 ([M + 2]⁺, 38), 242 (M⁺, 100), 211 (82), 183 (21).

9.5.1.12. 5-Etoxi-2-mercapto-1H-bencimidazol (**29**)

El producto se recristalizó de metanol-agua para dar un 86% de rendimiento de un polvo blanco que mostró ser un solo compuesto en ccf con R_f 0.55 (Sistema VI) y pf de 246.5–248.5 °C (lit: 245 °C) [101].

9.5.1.13. 6-Cloro-5-etoxi-2-mercapto-1H-bencimidazol (**30**)

El producto se recristalizó de acetona para dar un 78% de rendimiento de un sólido blanco que mostró ser un solo compuesto en ccf con R_f 0.55 (Sistema VI) y pf de 300.0–301.1 °C (dec). RMN ^1H (300 MHz, DMSO- d_6) δ : 1.338 (t, 3H, CH₃CH₂O–, J = 6.9 Hz), 4.056 (q, 2H, CH₃CH₂O–, J = 7.00 Hz), 6.813 y 7.149 (2s, 1H, H-4 y H-7), 12.472 y 12.580 (2s, 1H, SH y NH o 2NH); RMN ^{13}C (75.5 MHz, DMSO- d_6) δ : 14.62 (OCH₂CH₃), 64.91 (OCH₂CH₃), 95.46 (C-4), 110.37 y 116.25 (C-6 y C-7), 126.22 (C-3a), 131.88 (C-7a), 150.11 (C-5), 168.41 (C-2); EM, m/z (% intensidad relativa): 230 ([M + 2]⁺, 37), 228 (M⁺, 100), 202 (31), 201 (19), 200 (85), 199 (35).

9.5.1.14. 6-Etoxi-1-metil-2-mercapto-1H-bencimidazol (**31**)

El producto se recristalizó de etanol para dar un 85% de rendimiento de un sólido color beige que mostró ser un solo compuesto en ccf con R_f 0.66 (Sistema VII) y pf de 179.0–181.3 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 1.325 (t, 3H, CH₃CH₂O–, J = 6.9 Hz), 3.609 (s, 3H, NCH₃), 4.035 (q, 2H, CH₃CH₂O–, J = 7.0 Hz), 6.754 (dd, 1H, H-5, J = 8.7, J = 2.4 Hz), 6.981 (d, 1H, H-7, J = 2.1 Hz), 7.052 (d, 1H, H-4, J = 8.4 Hz),

12.572 (1H, SH o NH); RMN ^{13}C (75.5 MHz, DMSO- d_6) δ : 14.72 (OCH₂CH₃), 30.07 (NCH₃), 63.74 (OCH₂CH₃), 95.61 (C-7), 110.01 y 110.54 (C-4 y C-5), 124.63 (C-7a), 134.00 (C-3a), 155.04 (C-6), 168.18 (C-2); EM, m/z (% intensidad relativa): 208 (M^+ , 100), 180 (47), 179 (59), 151 (31).

9.5.1.15. 5-Cloro-6-etoxi-1-metil-2-mercapto-1H-bencimidazol (**32**)

El producto se recristalizó de metanol y se obtuvo un 80% de rendimiento de un polvo blanco que mostró ser un solo compuesto en ccf con R_f 0.58 (Sistema VII) y pf de 174.7–175.8 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 1.368 (t, 3H, CH₃CH₂O–, J = 6.9 Hz), 3.622 (s, 3H, NCH₃), 4.131 (q, 2H, CH₃CH₂O–, J = 6.90 Hz), 7.176 y 7.203 (2s, 1H, H-4 y H-7), 12.675 (s, 1H, SH o NH); RMN ^{13}C (75.5 MHz, DMSO- d_6) δ : 14.60 (OCH₂CH₃), 30.26 (NCH₃), 65.13 (OCH₂CH₃), 96.06 (C-7), 110.34 (C-5), 116.63 (C-4), 124.43 (C-7a), 132.80 (C-3a), 150.17 (C-6), 168.74 (C-2); EM, m/z (% intensidad relativa): 244 ($[\text{M} + 2]^+$, 34), 242 (M^+ , 100), 216 (34), 215 (30), 214 (96), 213 (61), 185 (44), 181 (49).

9.5.1.16. 5-Etoxi-1-metil-2-mercapto-1H-bencimidazol (**33**)

El producto se decoloró con carbón activado y se recristalizó de etanol-agua para dar un 81% de rendimiento de cristales color amarillo claro que mostraron ser un solo compuesto en ccf con R_f 0.56 (Sistema VII) y pf de 151.7–153.4 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 1.312 (t, 3H, CH₃CH₂O–, J = 6.9 Hz), 3.600 (s, 3H, NCH₃), 4.000 (q, 2H, CH₃CH₂O–, J = 6.9 Hz), 6.685 (d, 1H, H-4, J = 2.4 Hz), 6.780 (dd, 1H, H-6, J = 8.7, Hz, J = 2.4 Hz), 7.235 (d, 1H, H-7, J = 8.7 Hz), 12.602 (s, 1H, SH o NH); RMN ^{13}C (75.5 MHz, DMSO- d_6) δ : 14.70 (OCH₂CH₃), 30.05 (NCH₃), 63.64 (OCH₂CH₃), 95.07 (C-4), 110.01 y 110.06 (C-6 y C-7), 127.40 (C-7a), 131.41 (C-3a), 155.42 (C-5), 167.95 (C-2); EM, m/z (% intensidad relativa): 208 (M^+ , 100), 180 (44), 179 (74), 147 (37).

9.5.1.17. 6-Cloro-5-etoxi-1-metil-2-mercapto-1H-bencimidazol (**34**)

El producto se recristalizó de etanol-agua para dar un 69% de rendimiento como un polvo color hueso que mostró ser un solo compuesto en ccf con R_f 0.64 (Sistema XI) y con pf de 248.5–250.6 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 1.351 (t, 3H,

$\text{CH}_3\text{CH}_2\text{O}-$, $J = 6.8$ Hz), 3.587 (s, 3H, NCH_3), 4.071 (q, 2H, $\text{CH}_3\text{CH}_2\text{O}-$, $J = 6.7$ Hz), 6.825 (s, 1H, H-4), 7.521 (s, 1H, H-7), 12.758 (s, 1H, SH o NH); RMN ^{13}C (75.5 MHz, $\text{DMSO}-d_6$) δ : 14.58 (OCH_2CH_3), 30.20 (NCH_3), 65.00 (OCH_2CH_3), 95.30 (C-4), 110.74 (C-7), 116.67 (C-6), 127.26 (C-7a), 130.10 (C-3a), 150.437 (C-5), 168.58 (C-2); EM, m/z (% intensidad relativa): 244 ($[\text{M} + 2]^+$, 31), 242 (M^+ , 100), 216 (26), 215 (31), 214 (70), 213 (66), 185 (18), 181 (27).

9.6. Procedimiento de obtención de 1,2-Diaminobencenos (*o*-fenilendiaminas)

9.6.1. Procedimiento general de hidrogenación catalítica

Una mezcla del derivado de nitroanilina (0.037 mol) adecuadamente sustituida (**2–5** y **7–17**), 10% en peso de Pd/C al 5% para compuestos no clorados o 30% en peso de Ni-Raney para compuestos clorados, metanol (100 mL) y acetato de etilo (100 mL) se agitaron con cargas constantes de 60 lb de hidrógeno hasta consumirse la materia prima. El catalizador se eliminó por filtración al vacío y el filtrado se concentró a presión reducida dando la diamina correspondiente (**2–5** y **7–17**) que se usó inmediatamente en la siguiente reacción sin purificación.

9.6.2. 3,4-Diaminobenzoato de metilo (**6**)

Una mezcla de ácido 3,4-diaminobenzoico (7 g, 0.046 mol), metanol (70 mL) y H_2SO_4 concentrado (7 mL) se calentó a reflujo por 8 horas. La mezcla se vertió sobre agua helada y se neutralizó con carbonato de sodio. El precipitado formado se separó por filtración al vacío dando 7.2 g (91.0%) de agujas color café que mostraron ser un solo compuesto en ccf con R_f 0.32 (Sistema VI) y pf de 104.1–106.1 °C (lit: 104 °C) [97].

9.7. Preparación del 1-(2-Cloroetil)-1H-imidazol

9.7.1. 1-(2-Cloroetil)-1H-imidazol

Una mezcla de imidazol (5.1 g, 0.075 mol), KOH (28 g, 0.5 mol), K_2CO_3 (22.1 g, 0.16 mol), bromuro de tetrabutilamonio (0.52 g, 0.0016 mol) y 1,2-dicloroetano (200 mL), se calentó a 50 °C con agitación vigorosa por 3.5 horas. La fase orgánica se filtró

y se lavó con agua (3 x 30 mL), se secó con sulfato de sodio anhidro y se evaporó el disolvente a presión reducida. Se obtuvieron 7.4 g (75%) de un aceite amarillo que mostró una mancha principal en ccf (cámara de yodo) con R_f 0.56 (Sistema X). El producto crudo se usó como tal para la siguiente reacción [71].

9.8. Síntesis de precursores

9.8.1. Procedimientos generales

9.8.1.1. Procedimiento general de acetilación

El derivado de anilina adecuadamente sustituido se agregó lentamente sobre anhídrido acético (2–4 eq) manteniendo la temperatura por debajo de 70 °C. Para los compuestos con un grupo nitro en posición *orto* con respecto a la amina se agregó ácido sulfúrico concentrado como catalizador (3–5 gotas). La mezcla de reacción se mantuvo en agitación a 70 °C durante 30 minutos y se vertió sobre agua con hielo (200–500 mL). El sólido formado se separó por filtración al vacío, se lavó varias veces con agua hasta pH neutro y se dejó secar.

9.8.1.2. Procedimiento general para hidrólisis de acetamidas en ácido sulfúrico concentrado

El derivado de acetamida correspondiente se disolvió en ácido sulfúrico concentrado (cantidad mínima necesaria para disolver el compuesto) y se agregó agua cuidadosamente (1 eq). La mezcla se llevó a 80 °C durante 30 minutos, se dejó enfriar y se vertió sobre hielo (200–500 g). El sólido formado se separó por filtración al vacío, se lavó con agua helada hasta pH neutro y se dejó secar.

9.8.1.3. Procedimiento general para hidrólisis de acetamidas en ácido sulfúrico y metanol

La acetamida adecuadamente sustituida se disolvió (1/10 m/v) en ácido sulfúrico al 10% en metanol. A la mezcla de reacción se le agregó agua (1 eq) y se calentó a reflujo durante 30 minutos. Una vez consumida la materia prima, la reacción

se vertió sobre hielo (500 g) y el sólido formado se separó por filtración al vacío, se lavó con agua helada hasta pH neutro y se dejó secar.

9.8.1.4. Procedimiento general para hidrólisis de acetamidas en medio básico

Una mezcla de la acetamida adecuadamente sustituida, metanol (1/10 m/v), KOH (1.1 eq) y agua (1 eq), se llevó temperatura entre 25–60 °C. Una vez consumida la materia prima, la reacción se dejó enfriar y se vertió sobre hielo (500 g). El sólido formado se separó por filtración al vacío, se lavó con agua helada hasta pH neutro y se dejó secar.

9.8.1.5. Procedimiento general de N-metilación de acetamidas

A una solución de acetamida adecuadamente sustituida en monoglina (1/10 a 1/50 m/v) y sulfato de dimetilo (1.2 eq) se añadió lentamente hidróxido de sodio al 50% (1.5 eq) manteniendo la temperatura entre 30–35 °C. Terminada la adición, la reacción se dejó en agitación por una hora más. La reacción se vertió sobre hielo (500 g) formándose un precipitado que se separó por filtración al vacío y se lavó con agua helada.

9.8.1.6. Procedimiento general para la preparación de ésteres con sulfato de dimetilo

A una mezcla del derivado de ácido benzoico adecuadamente sustituido, DMF (proporción 1/5 m/v) y agua (proporción 1/0.5 m/v), se agregó NaHCO₃ (1.1 eq). Después se adicionó lentamente sulfato de dimetilo (1.5 eq) y se calentó a 65 °C por 2 horas. Una vez agotada la materia prima, la mezcla se vertió sobre agua helada (500 mL) para formar un precipitado que se separó por filtración al vacío, se lavó con agua y se dejó secar.

9.8.1.7. Procedimiento general para realizar la esterificación de Fischer

El ácido benzoico adecuadamente sustituido se disolvió (1/20 m/v) en ácido sulfúrico al 10% en metanol. La mezcla de reacción se calentó a reflujo hasta agotarse la materia prima (3–8 horas). Posteriormente, la mezcla se vertió sobre agua helada

(500 mL) y se neutralizó con carbonato de sodio al 10 % para formar un precipitado que se separó por filtración al vacío, se lavó con agua y se dejó secar.

9.8.1.8. Procedimiento general de alquilación de fenoles

A una mezcla de fenol adecuadamente sustituido, DMF (1/10 m/v), agua (1/1 m/v), se agregó carbonato de potasio (1.2 eq). Posteriormente se agregó yoduro de etilo (1.3 eq) y se calentó a 50 °C durante toda la noche. Una vez terminada la materia prima, la reacción se vertió en agua helada (300 mL) y el precipitado formado se separó por filtración al vacío, se lavó con solución de bicarbonato de sodio al 5% (100 mL) y se dejó secar.

9.8.2. Preparación de intermediarios

9.8.2.1. *N*-(3,4-Diclorofenil)acetamida (**Int-3**)

Siguiendo el procedimiento general de acetilación (sección 9.8.1.1), partiendo de 3,4-dicloroanilina (50 g, 0.3086 mol), se obtuvieron 62.6 g (99%) de **Int-3** como un sólido blanco que mostró ser un solo compuesto en ccf con R_f 0.6 (Sistema VI). Una muestra se recrystalizó de etanol dando cristales blancos con pf de 124.7–126.4 °C (lit: 121.9–122.3 °C) [102].

9.8.2.2. *N*-(4,5-Dicloro-2-nitrofenil)acetamida (**Int-4**)

La *N*-(4,5-diclorofenil)acetamida (50 g, 0.2450 mol) se disolvió en H₂SO₄ concentrado (200 mL) manteniendo la temperatura debajo de 10 °C; posteriormente se añadió lentamente y con agitación constante HNO₃ (20.5 mL, 0.2959 mol, 1.21 eq) manteniendo la temperatura debajo de 5 °C. Después de la adición, la reacción se dejó en agitación por 20 minutos adicionales y se vertió sobre hielo-agua (1 kg). El precipitado formado se separó por filtración al vacío y se lavó con agua varias veces. El producto crudo se suspendió sobre metanol (250 mL) y se dejó en agitación durante toda la noche a temperatura ambiente; posteriormente, el sólido se separó por filtración al vacío lavando con metanol (2 x 50 mL) y se dejó secar. Se obtuvieron 49.39 g (81%) del **Int-4** como un sólido amarillo que mostró ser un solo compuesto en

ccf con R_f 0.62 (Sistema I). Una muestra se recrystalizó de metanol dando cristales amarillos con pf de 127.0–129.2 °C (lit: 122.0–124.0 °C) [103].

9.8.2.3. 4,5-Dicloro-2-nitroanilina (**Int-5**)

Siguiendo el procedimiento general de hidrólisis (sección 9.8.1.2), partiendo de *N*-(4,5-dicloro-2-nitrofenil)acetamida (40 g, 0.1606 mol), se obtuvieron 32.3 g (97%) de **Int-5** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.74 (Sistema XII). Una muestra se recrystalizó de ácido acético dando cristales amarillos con pf de 178.7–179.1 °C (lit: 178.0–181.0 °C) [104].

9.8.2.4. 2,4-Dicloronitrobenceno (**Int-7**)

Sobre una mezcla con agitación de 1,3-diclorobenceno (40 mL, 0.3504 mol) y H₂SO₄ concentrado (200 mL) se añadió lentamente HNO₃ al 30% (57 mL, 0.3848 mol, 1.1 eq) manteniendo la temperatura de –5 a 5 °C. Después, la reacción se dejó en agitación por 20 minutos adicionales y se vertió sobre agua helada (500 mL); posteriormente, el agua se decantó y el producto se neutralizó con una solución al 5% de bicarbonato de sodio. El producto crudo se extrajo con acetato de etilo (3 x 100 mL), los extractos orgánicos reunidos se lavaron con agua fría, se secaron con Na₂SO₄ anhidro y el disolvente se evaporó a presión reducida. El producto obtenido se recrystalizó de hexano dando 54.75 g (83%) de **Int-7** como cristales en forma de agujas color amarillo pálido que mostraron ser un solo compuesto en ccf con R_f 0.48 (Sistema XIII) y pf de 30.0–31.5 °C (lit: 31.5–33.0 °C) [95].

9.8.2.5. 5-Cloro-*N*-metil-2-nitroanilina (**Int-8**)

Una mezcla de 2,4-dicloronitrobenceno (10.01 g, 0.0521 mol), K₂CO₃ (36 g, 0.2605 mol, 5 eq), CH₃NH₂.HCl (17.75 g, 0.2629 mol, 5 eq), DMSO (80 mL) y agua (20 mL), se calentó a 60 °C durante 16 horas. Posteriormente, la mezcla de reacción se vertió en hielo-agua (500 g) y el precipitado se separó por filtración al vacío. El producto crudo se recrystalizó de metanol obteniendo 7.89 g (81%) de **Int-8** como

cristales en forma de agujas color amarillo que mostraron ser un solo compuesto en ccf con R_f 0.26 (Sistema XIV) y pf de 105.9–106.6 °C (lit: 100.0–102.0 °C) [105].

9.8.2.6. *N*-(4-Cloro-2-nitrofenil)acetamida (**Int-9**)

Siguiendo el procedimiento general de acetilación (sección 9.8.1.1), partiendo de 4-cloro-2-nitroanilina (25 g, 0.1449 mol), se obtuvieron 30.5 g (98%) de **Int-9** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.68 (Sistema VI). Una muestra se recristalizó de etanol dando cristales amarillos con pf de 99.6–101.1 °C (lit: 98–100 °C) [106].

9.8.2.7. *N*-(4-Cloro-2-nitrofenil)-*N*-metilacetamida (**Int-10**)

Siguiendo el procedimiento general de *N*-metilación de acetamidas (sección 9.8.1.5), partiendo de *N*-(4-cloro-2-nitrofenil)acetamida (30 g, 0.1398 mol), se obtuvieron 29.61 g (93%) de **Int-10** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.58 (Sistema VII). Una muestra se recristalizó de etanol-agua para dar cristales amarillos con pf de 99.6–101.1 °C (lit: 98–100 °C) [106].

9.8.2.8. 4-Cloro-*N*-metil-2-nitroanilina (**Int-11**)

Siguiendo el procedimiento general de hidrólisis (sección 9.8.1.2), partiendo de *N*-(4-cloro-2-nitrofenil)-*N*-metilacetamida (27 g, 0.1181 mol), se obtuvieron 18.76 g (85%) un polvo rojo ladrillo que mostró una sola mancha en ccf con R_f 0.52 (Sistema V). Una muestra se recristalizó de etanol-agua para dar cristales rojo ladrillo en forma de agujas con pf de 108.9–110.0 °C (lit: 107–109 °C) [107].

9.8.2.9. 3-Metoxi-4-nitrobenzoato de metilo (**Int-15**)

Una mezcla de ácido 3-hidroxi-4-nitrobenzoico (5 g, 0.0273 mol), K_2CO_3 (11.31 g, 0.0818 mol, 3 eq), acetona (40 mL) y agua (20 mL) se agitó a temperatura ambiente durante 10 min. Posteriormente se agregó sulfato de dimetilo (7.8 mL, 0.0819 mol, 3 eq) y la mezcla se calentó a 40 °C por 4 horas. Una vez agotada la materia prima, la mezcla se vertió sobre agua helada (300 mL) formando un precipitado que se separó

por filtración al vacío. Se obtuvieron 5.2 g (90%) de **Int-15** como un sólido amarillo que mostró ser un solo compuesto en ccf con un R_f 0.76 (Sistema I). Una muestra se recrystalizó de MeOH-agua dando cristales en forma de agujas amarillas con pf de 91.0–92.2 °C [108].

9.8.2.10. *Ácido 3-metoxi-4-nitrobenzoico (Int-16)*

Siguiendo el procedimiento general de hidrólisis (sección 9.8.1.4), partiendo de 3-metoxi-4-nitrobenzoato de metilo (6 g, 0.0284 mol), se obtuvieron 5.49 g (98%) de **Int-16** como un producto amarillo que mostró una sola mancha en ccf con R_f 0.35 (Sistema I). Una muestra se recrystalizó de metanol-agua para dar agujas amarillas con punto de fusión de 232.8–234.2 °C (lit: 230–233 °C) [109].

9.8.2.11. *Ácido 3-(metilamino)-4-nitrobenzoico (Int-17)*

Una mezcla de ácido 3-metoxi-4-nitrobenzoico (10.63 g, 0.0539 mol), carbonato de potasio (29.8 g, 0.2156 mol, 4 eq), clorhidrato de metilamina (10.92 g, 0.1617 mol, 3 eq) y agua (150 mL) se calentó en un reactor Parr a 105 °C por 5 días. Posteriormente, la mezcla se vertió sobre agua helada (500 mL) y se acidificó con ácido clorhídrico para dar un precipitado naranja. El producto se separó por filtración al vacío, se lavó con agua y se dejó secar. Se obtuvieron 10.47 g (99%) de un sólido anaranjado que mostró ser un solo compuesto en ccf con R_f 0.71 (XV) y mostró un punto de fusión de 272.0–272.8 °C [110].

9.8.2.12. *3-(Metilamino)-4-nitrobenzoato de metilo (Int-18)*

Siguiendo el procedimiento general de esterificación (sección 9.8.1.7), partiendo de ácido 3-(metilamino)-4-nitrobenzoico (9.25 g, 0.0467 mol), se obtuvieron 8.64 g (88%) de un sólido anaranjado que mostró ser un solo compuesto en ccf con R_f 0.57 (Sistema I) y con un punto de fusión de 134.9–135.8 °C.

9.8.2.13. 2-Cloro-5-nitrobenzoato de metilo (**Int-20**)

Siguiendo el procedimiento general de esterificación (sección 9.8.1.6), partiendo de ácido 2-cloro-5-nitrobenzoico (40 g, 0.1985 mol), se obtuvieron 42.35 g (99%) de un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.80 (Sistema VI). Una muestra se recrystalizó de etanol para dar cristales amarillos con punto de fusión de 61.7–64.5 (lit: 67–68 °C) [111] (68–69 °C) [112] (60–63 °C) [113]

9.8.2.14. 5-Amino-2-clorobenzoato de metilo (**Int-21**)

Siguiendo el procedimiento general de hidrogenación catalítica (sección 9.6.1), partiendo de 2-cloro-5-nitrobenzoato de metilo (15 g, 0.0696 mol), se obtuvieron 12.9 g (100%) de un sólido verde que mostró una sola mancha en ccf y R_f 0.75 (Sistema IV). El producto crudo se usó como tal para la siguiente reacción.

9.8.2.15. 5-(Acetamino)-2-clorobenzoato de metilo (**Int-22**)

Siguiendo el procedimiento general de acetilación (sección 9.8.1.1), partiendo de 5-amino-2-clorobenzoato de metilo (30 g, 0.1616 mol), se obtuvieron 34.9 g (95%) de **Int-22** como un sólido blanco que mostró una sola mancha en ccf con R_f 0.38 (Sistema XVI). Una muestra se recrystalizó de tolueno dando cristales blancos con punto de fusión de 95.0–96.0 °C. RMN ^1H (400 MHz, CDCl_3) δ : 2.180 (s, 3H, –COCH₃), 3.898 (s, 3H, –COOCH₃), 7.362 (d, 1H, 9.2 Hz), 7.679 (dd, 1H, J = 2.6 y 9.0 Hz), 7.955 (d, 1H, 2.0 Hz), 8.022 (s, NH).

9.8.2.16. 5-(Acetamino)-2-cloro-4-nitrobenzoato de metilo (**Int-23a**)

La preparación del intermediario **Int-23a** se realizó mediante una secuencia de tres reacciones consecutivas.

Primero, el 5-(acetamino)-2-clorobenzoato de metilo (30 g, 0.1318 mol) se disolvió en H_2SO_4 concentrado (120 mL) manteniendo la temperatura de 5 a 10 °C, luego se añadió lentamente, con agitación constante, HNO_3 (10.05 mL, 0.1448 mol, 1.1 eq) manteniendo la temperatura entre 0–5 °C; posteriormente, la reacción se dejó en agitación por 20 minutos adicionales y se vertió sobre hielo (1 kg). El precipitado

formado se separó por filtración al vacío y se dejó secar. Se obtuvieron 33.8 g (94%) de un producto amarillo que mostró dos machas en ccf en proporciones iguales con R_f 0.40 (**Int-23b**) y 0.51 (**Int-23a**), respectivamente (Sistema XVI). Una muestra analítica se separó usando cromatografía en columna dando el producto **Int-23a** y **Int-23b**. **Int-23a**: RMN ^1H (300 MHz, CDCl_3) δ : 2.315 (s, 3H, $-\text{COCH}_3$), 3.976 (s, 3H, $-\text{COOCH}_3$), 8.287 (s, 1H), 9.210 (s, 1H), 10.168 (s, 1H, NH); RMN ^{13}C (75 MHz, CDCl_3) δ : 25.513, 53.142, 124.514, 126.900, 127.376, 133.057, 136.883, 164.347, 168.880. **Int-23b**: RMN ^1H (300 MHz, CDCl_3) δ : 2.279 (s, 3H, $-\text{COCH}_3$), 3.977 (s, 3H, $-\text{COOCH}_3$), 7.638 (dd, 1H, $J = 9.2$ Hz, 0.45 Hz), 8.655 (d, 1H, $J = 9$ Hz), 9.557 (s, 1H, NH); RMN ^{13}C (75 MHz, CDCl_3) δ : 25.306, 53.612, 124.643, 126.815, 130.036, 132.447, 135.498, 163.901, 168.801.

La mezcla de productos **Int-23a** y **Int-23b** (30g, 0.1100 mol) se hidrolizó siguiendo la metodología general descrita en la sección 9.8.1.3. Una vez terminada la reacción el producto cristalizó y se separó por filtración al vacío para dar 6.6 g (26%) de **Int-23c** que mostró ser un solo compuesto en ccf con R_f 0.56 (Sistema XVI) y pf de 181.6–184.1 (lit: 225–226 °C) [114].

Finalmente, siguiendo el procedimiento general de acetilación (sección 9.8.1.1), partiendo de **Int-23c** (6.1 g, 0.0265 mol), se obtuvieron 7.3 g (100%) de **Int-23a** como un polvo amarillo que mostró ser un solo compuesto en ccf con R_f 0.51 (Sistema XVII). Una muestra se recrystalizó de metanol para dar cristales amarillos con pf de 115.7–116.6 °C (referencia obtenida de la separación cromatográfica de productos de nitración 115–117 °C).

El rendimiento global de los tres pasos descritos en esta sección fue de 24%.

9.8.2.17. 5-[Acetil(metil)amino]-2-cloro-4-nitrobenzoato de metilo (**Int-24**)

Siguiendo el procedimiento general de metilación (sección 9.8.1.5), partiendo de 5-(acetamino)-2-cloro-4-nitrobenzoato de metilo (5.7 g, 0.0211 mol), se obtuvieron 5.1 g (77%) de **Int-24** como un sólido amarillo que mostró ser un solo compuesto por ccf con R_f 0.23 (Sistema II). Una muestra se recrystalizó de hexano-acetato de etilo para dar cristales amarillos con pf de 116.7–118.7 °C.

9.8.2.18. 2-Cloro-5-(metilamino)-4-nitrobenzoato de metilo (**Int-25**)

Siguiendo el procedimiento general de hidrólisis (sección 9.8.1.3), partiendo de 5-[acetil(metil)amino]-2-cloro-4-nitrobenzoato de metilo (5 g, 0.0174 mol), se obtuvieron 4.06 g (88%) de **Int-25** como cristales color rojo ladrillo en forma de agujas que mostraron ser un solo compuesto en ccf con R_f 0.7 (Sistema I) y pf de 148.7–150.8 °C. RMN ^1H (300 MHz, DMSO- d_6) δ : 2.936 (d, 3H, $-\text{NHCH}_3$, $J = 5.1$ Hz), 3.864 (s, 3H, $-\text{COOCH}_3$), 7.279 (s, 1H), 8.102 (s, 1H), 8.242 (d, NH, 3.9 Hz); RMN ^{13}C (75 MHz, DMSO- d_6) δ : 29.951, 53.113, 114.917, 116.632, 127.133, 131.642, 137.555, 144.200, 165.092.

9.8.2.19. 4-Cloro-3-nitrobenzoato de metilo (**Int-27**)

Siguiendo el procedimiento general de esterificación (sección 9.8.1.6), partiendo de ácido-4-cloro-3-nitrobenzoico (30 g, 0.1488 mol), se obtuvieron 31.33 g (98%) de **Int-27** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f de 0.41 (Sistema I). Una muestra se recrystalizó de metanol dando cristales amarillos con punto de fusión de 82.0–87.3 °C (lit: 83 °C) [115].

9.8.2.20. 4-(Metilamino)-3-nitrobenzoato de metilo (**Int-28**)

Una mezcla de 4-cloro-3-nitrobenzoato de metilo (15 g, 0.0696 mol), carbonato de potasio (47.99 g, 0.3473 mol, 5 eq), clorhidrato de metilamina (18.79 g, 0.2782 mol, 4 eq), agua (10 mL) y monoglima (417 mL) se calentó a 130 °C en un reactor Parr por 2.5 horas. Posteriormente, la mezcla de reacción se vertió sobre agua helada (1 L) y se acidificó con ácido clorhídrico al 10%. El precipitado formado se separó por filtración al vacío y se dejó secar dando 13.78 g (94%) de **Int-28** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.44 (Sistema I). Una muestra se recrystalizó de etanol y mostró un punto de fusión de 145.2–146.0 °C (lit: 142–143 °C) [116] (lit: 145 °C) [117]

9.8.2.21. 4-[Acetil(metil)amino]-2-cloro-5-nitrobenzoato de metilo (**Int-30**)

Siguiendo el procedimiento general de metilación (sección 9.8.1.5), partiendo de 4-(acetamino)-2-cloro-5-nitrobenzoato de metilo (10 g, 0.0367 mol), se obtuvieron 9.7 g (92%) de **Int-30** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.16 (Sistema I). Una muestra se recrystalizó de hexano-acetato de etilo para dar cristales en forma de agujas amarillas con pf de 89.4–91.1 °C.

9.8.2.22. 2-Cloro-4-(metilamino)-5-nitrobenzoato de metilo (**Int-31**)

Siguiendo el procedimiento general de hidrólisis (sección 9.8.1.3), partiendo de 4-[acetil(metil)amino]-2-cloro-5-nitrobenzoato de metilo (6 g, 0.0209 mol), se obtuvieron 5 g (98%) de **Int-31** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.60 (Sistema I). Una muestra se recrystalizó de etanol con unas gotas de DMF para dar cristales amarillos en forma de agujas con pf de 191.2–192.5 °C.

9.8.2.23. 1-Etoxi-4-nitrobenceno (**Int-33**)

Siguiendo el procedimiento general de alquilación (sección 9.8.1.8), para 25 g (0.1797 mol) de 4-nitrofenol se obtuvieron 28.14 g (94%) de **Int-33** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.42 (Sistema IV). Una muestra se recrystalizó de etanol dando cristales amarillos con punto de fusión de 58.0–59.5 °C (lit: 56–60 °C) [118].

9.8.2.24. 4-Etoxianilina (**Int-34**)

Siguiendo el procedimiento general de hidrogenación catalítica (sección 9.6.1), partiendo de 4-etoxinitrobenceno (27.5 g, 0.1645 mol), se obtuvieron 22.56 g (100%) de **Int-34** como un aceite café-rojizo que mostró una sola mancha en ccf con R_f 0.63 (Sistema X). El producto crudo se usó como tal para la siguiente reacción.

9.8.2.25. *N*-4-(Etoxifenil)acetamida (**Int-35**)

Siguiendo el procedimiento general de acetilación (sección 9.8.1.1), partiendo de 4-etoxianilina (21 g, 0.1531 mol), se obtuvieron 26.3 g (96%) de **Int-35** como un sólido gris claro que mostró ser un solo compuesto en ccf con R_f 0.33 (Sistema VII). Una muestra se recrystalizó de etanol dando cristales blancos en forma de agujas con punto de fusión 134.5–135.5 °C (lit: 135 °C) [119].

9.8.2.26. *N*-(4-Etoxi-2-nitrofenil)acetamida (**Int-36**)

A una mezcla de *N*-(4-etoxifenil)acetamida (25 g, 0.1394 mol), agua (50 mL) y ácido acético (38.5 mL) se añadió lentamente ácido nítrico (15.5 mL, 0.2237 mol, 1.6 eq) conservando la temperatura a 10 °C. Posteriormente, la reacción se calentó a 40 °C por una hora. Una vez agotada la materia prima, la mezcla de reacción se vertió sobre hielo-agua (1 kg) y el precipitado formado se separó por filtración al vacío y se dejó secar. Finalmente, el producto se purificó por recrystalización de etanol para dar 24.2 g (77%) de cristales amarillos en forma de agujas que mostraron ser un solo compuesto en ccf con R_f 0.40 (Sistema I) y pf de 102.3–103.4 °C (lit: 101–103 °C) [120–121].

9.8.2.27. 4-Etoxi-2-nitroanilina (**Int-37**)

Siguiendo el procedimiento general de hidrólisis (sección 9.8.1.4), partiendo de *N*-(4-etoxi-2-nitrofenil)acetamida (10 g, 0.0111 mol), se obtuvieron 8.04 g (99%) de **Int-37** como cristales anaranjado-rojizos que mostraron ser un solo compuesto en ccf con R_f 0.49 (Sistema I) y pf de 111.6–112.9 °C (lit: 110–112 °C) [121].

9.8.2.28. 4-Cloro-5-etoxi-2-nitroanilina (**Int-38**)

Sobre etanol anhidro (100 mL), con agitación constante y en baño de hielo, se añadió poco a poco sodio metálico (2.5 g, 0.1109 mol, 1.5 eq). Una vez consumido el sodio se añadió 4,5-dicloro-2-nitroanilina (15 g, 0.0725 mol) y se calentó a reflujo durante 6 horas. Posteriormente, la reacción se enfrió y se vertió en agua helada (500 mL); el precipitado formado se filtró al vacío para dar 13.8 g (88%) de un sólido

amarillo intenso el cual mostró una sola macha en cromatografía en capa fina con R_f 0.59 (Sistema XI). Una muestra se recrystalizó de metanol dando cristales en forma de agujas amarillas con punto de fusión de 149.4–151.1 °C [122].

9.8.2.29. 3-Cloro-4-nitrofenol (**Int-40**)

Sobre una mezcla de 3-clorofenol (30 g, 0.2334 mol) y ácido acético glacial (30 mL) se añadió, lentamente y manteniendo la temperatura entre –5 y 0 °C, ácido nítrico (19 mL, 0.2742 mol, 1.2 eq) diluido en ácido acético glacial (53 mL). Posteriormente, la mezcla de reacción se dejó en agitación por una hora más. Una vez agotada la materia prima, el producto se vertió sobre agua helada (200 mL) y se extrajo con acetato de etilo (3 x 100 mL). La fase orgánica se lavó con agua (100 mL), se secó con sulfato de sodio anhidro y se evaporó en el rotavapor. El residuo se trató con benceno (50 mL) y se filtró al vacío dando un producto color anaranjado que se recrystalizó de benceno para dar 11.82 g (29%) de cristales amarillos en forma de agujas que mostraron ser un solo compuesto por ccf con R_f 0.35 (Sistema I) y pf de 120.6–122.9 °C [123].

9.8.2.30. 2-Cloro-4-etoxi-1-nitrobenzoceno (**Int-41**)

Siguiendo el procedimiento general de alquilación (sección 9.8.1.8), partiendo de 3-cloro-4-nitrofenol (10 g, 0.0576 mol), se obtuvieron 10.9 g (94%) de **Int-41** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.67 (Sistema III). Una muestra se recrystalizó de metanol-agua dando un sólido color amarillo pálido con pf de 37.7–38.5 °C (lit: 40 °C) [124].

9.8.2.31. 5-Etoxi-N-metil-2-nitroanilina (**Int-42**)

Una mezcla de 2-cloro-4-etoxi-1-nitrobenzoceno (10 g, 0.0496 mol), carbonato de potasio (27.5 g, 0.1990 mol, 4 eq), clorhidrato de metilamina (10 g, 0.1481 mol, 3 eq), agua (15 mL) y metanol (70 mL) se calentó a 110 °C por 40 horas en un reactor Parr. Una vez agotada la materia prima, la mezcla de reacción se vertió sobre agua helada (300 mL) y se neutralizó con ácido clorhídrico al 10%. El precipitado formado se separó por filtración al vacío dando 7.2 g (74%) de **Int-42** como un sólido amarillo que

mostró ser un solo compuesto en ccf con R_f 0.67 (Sistema I). Una muestra se recrystalizó de metanol dando agujas amarillas con pf de 110.5–111.5 °C.

9.8.2.32. *N*-(4-Cloro-5-etoxi-2-nitrofenil)acetamida (**Int-43**)

Siguiendo el procedimiento general de acetilación (sección 9.8.1.1), partiendo de 4-cloro-5-etoxi-2-nitroanilina (7 g, 0.0355 mol), se obtuvieron 8.7 g (95%) de **Int-43** como un sólido amarillo pálido que mostró ser un solo compuesto en ccf con R_f 0.63 (Sistema III). Una muestra se recrystalizó de etanol dando cristales amarillos en forma de agujas con pf de 130.0–132.2 °C.

9.8.2.33. *N*-(4-Cloro-5-etoxi-2-nitrofenil)-*N*-metilacetamida (**Int-44**)

Siguiendo el procedimiento general de metilación (sección 9.8.1.5), partiendo de *N*-(4-cloro-5-etoxi-2-nitrofenil)acetamida (7 g, 0.0271 mol), se obtuvieron 7.06 g (96%) de **Int-44** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.58 (Sistema VI). Una muestra se recrystalizó de hexano-acetato de etilo para dar cristales amarillos con pf de 114.1–117.0 °C.

9.8.2.34. 4-Cloro-5-etoxi-*N*-metil-2-nitroanilina (**Int-45**)

Siguiendo el procedimiento general de hidrólisis (sección 9.8.1.3), partiendo de *N*-(4-cloro-5-etoxi-2-nitrofenil)-*N*-metilacetamida (6.6 g, 0.0242 mol), se obtuvieron 4.86 g (87%) de **Int-49** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.58 (Sistema I) y con pf de 172.0–174.4 °C.

9.8.2.35. *N*-(4-Etoxi-2-nitrofenil)-*N*-metilacetamida (**Int-46**)

Siguiendo el procedimiento general de metilación (sección 9.8.1.5), partiendo de *N*-(4-etoxi-2-nitrofenil)acetamida (8 g, 0.0357 mol), se obtuvieron 8.02 g (94%) de **Int-46** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.60 (Sistema VII). Una muestra se recrystalizó de hexano-acetato de etilo para dar cristales amarillos con pf de 165.2–167.2 °C.

9.8.2.36. 4-Etoxi-N-metil-2-nitroanilina (**Int-47**)

Siguiendo el procedimiento general de hidrólisis (sección 9.8.1.4), partiendo de *N*-(4-etoxi-2-nitrofenil)-*N*-metilacetamida (7.5 g, 0.0315 mol), se obtuvieron 5.88 g (95%) de **Int-47** como cristales anaranjado-rojizos que mostraron ser un solo compuesto en ccf con R_f 0.7 (Sistema I) con pf de 78.8–80.1 °C (lit: 78.5–79.0 °C) [125].

9.8.2.37. 2-Cloro-1-etoxi-4-nitrobenceno (**Int-49**)

Sobre etanol anhidro (300 mL), con agitación constante y baño de hielo, se añadió poco a poco sodio metálico (5.1 g, 0.2217 mol, 1.3 eq). Una vez consumido el sodio se añadió 2-cloro-1-fluoro-4-nitrobenceno (30 g, 0.1709 mol) manteniendo la temperatura por debajo de 40 °C. Posteriormente, la mezcla se dejó en agitación por 30 minutos y se vertió sobre agua (500 mL). El precipitado formado se separó por filtración al vacío obteniéndose 34.4 g (100%) de un sólido amarillo intenso que mostró ser un solo compuesto en ccf con R_f 0.42 (Sistema XIV). Una muestra se recrystalizó de metanol dando cristales en forma de agujas blancas con pf de 82.0–83.5 °C (lit: 81.5–83.0) [122].

9.8.2.38. 3-Cloro-4-etoxianilina (**Int-50**)

Siguiendo el procedimiento general de hidrogenación catalítica (sección 9.6.1), partiendo de 2-cloro-1-etoxi-4-nitrobenceno (17.2 g, 0.0853 mol), se obtuvieron 14.6 g (100%) de **Int-50** como un sólido rosa claro que mostró una sola mancha en ccf y un R_f 0.63 (Sistema VIII). Una muestra se recrystalizó de éter de petróleo para dar cristales con pf de 64.0–66.0 °C (lit: 64–65 °C) [126].

9.8.2.39. *N*-(3-Cloro-4-etoxifenil)acetamida (**Int-51**)

Siguiendo el procedimiento general de acetilación (sección 9.8.1.1), partiendo de 3-cloro-4-etoxianilina (14.6 g, 0.0851 mol), se obtuvieron 17.1 g (94%) de **Int-51** como un sólido blanco que mostró ser un solo compuesto en ccf con R_f 0.34 (Sistema

VI). Una muestra se recrystalizó de etanol para dar cristales blancos en forma de agujas con pf de 132.2–134.3 °C (lit: 128–129 °C) [126].

9.8.2.40. *N*-(5-Cloro-4-etoxi-2-nitrofenil)acetamida (**Int-52**)

A una solución de *N*-(3-cloro-4-etoxifenil)acetamida (10 g, 0.0468 mol) en ácido acético (50 mL) se agregó lentamente y con agitación constante ácido nítrico (3.56 mL, 0.0514 mol, 1.1 eq) diluido en ácido sulfúrico (2.7 mL) manteniendo la temperatura por debajo de 30 °C. Una vez agotada la materia prima, la mezcla de reacción se vertió sobre hielo-agua (500 g) y el precipitado formado se separó por filtración al vacío. Se obtuvieron 11.89 g (98%) de **Int-52** como un sólido amarillo que mostró ser un solo compuesto por cromatografía en capa fina con R_f 0.5 (Sistema I). Una muestra se recrystalizó de etanol-cloroformo para dar cristales en forma de agujas amarillas con pf de 166.1–168.6 °C (lit: 166 °C) [127].

9.8.2.41. 5-Cloro-4-etoxi-2-nitroanilina (**Int-53**)

Siguiendo el procedimiento general de hidrólisis (sección 9.8.1.4), partiendo de *N*-(5-Cloro-4-etoxi-2-nitrofenil)acetamida (10 g, 0.0387 mol), se obtuvieron 8.07 g (96%) de **Int-47** como cristales anaranjado-rojizos que mostraron ser un solo compuesto en ccf con R_f 0.56 (Sistema I) con pf de 131.6-134.0 °C (lit: 128.5) [127].

9.8.2.42. *N*-(5-Cloro-4-etoxi-2-nitrofenil)-*N*-metilacetamida (**Int-54**)

Siguiendo el procedimiento general de metilación (sección 9.8.1.5), partiendo de *N*-(5-cloro-4-etoxi-2-nitrofenil)acetamida (8 g, 0.0309 mol), se obtuvieron 7.61 g (90%) de **Int-54** como un sólido amarillo que mostró ser un solo compuesto en ccf con R_f 0.33 (Sistema I). Una muestra se recrystalizó de hexano-acetato de etilo para dar cristales amarillos con pf de 107.0–108.2 °C.

9.8.2.43. 5-Cloro-4-etoxi-*N*-metil-2-nitroanilina (**Int-55**)

Siguiendo el procedimiento general de hidrólisis (sección 9.8.1.4), partiendo de *N*-(5-cloro-4-etoxi-2-nitrofenil)-*N*-metilacetamida (7g, 0.0256 mol), se obtuvieron 4.6 g (78%) de **Int-55** como cristales rojos que mostraron ser un solo compuesto en ccf con R_f de 0.66 (Sistema I) y pf de 142.5–145.0 °C.

10. REFERENCIAS

- [1] Boletín epidemiología 2012, http://www.dgepi.salud.gob.mx/2010/plantilla/intd_boletin.html, 2012.
- [2] P. Upcroft, J.A. Upcroft, Drug targets and mechanisms of resistance in the anaerobic protozoa, *Clinical Microbiology Reviews*, 14 (2001) 150–164.
- [3] D.S. Berkman, A.G. Lescano, R.H. Gilman, S. Lopez, M.M. Black, Effects of stunting, diarrhoeal disease, and parasitic infection during infancy on cognition in late childhood: a follow-up study, *Lancet*, 359 (2002) 564–571.
- [4] M. Tanyuksel, W.A. Petri, Laboratory diagnosis of amebiasis, *Clinical Microbiology Reviews*, 16 (2003) 713–729.
- [5] S.L. Cudmore, K.L. Delgaty, S.F. Hayward-McClelland, D.P. Petrin, G.E. Garber, Treatment of infections caused by metronidazole-resistant *Trichomonas vaginalis*, *Clinical Microbiology Reviews*, 17 (2004) 783–793.
- [6] Global prevalence and incidence of selected curable sexually transmitted infections: overview and estimates, World Health Organization, Geneva, (2001).
- [7] S.D. Sears, J. Ohare, In vitro susceptibility of *Trichomonas vaginalis* to 50 antimicrobial agents, *Antimicrobial Agents and Chemotherapy*, 32 (1988) 144–146.
- [8] R. Cedillo-Rivera, O. Munoz, In vitro susceptibility of *Giardia lamblia* to albendazole, mebendazole and other chemotherapeutic agents, *Journal of Medical Microbiology*, 37 (1992) 221–224.
- [9] B. Chavez, R. Cedillo-Rivera, A. Martinez-Palomo, *Giardia lamblia* ultrastructural study of the in vitro effect of benzimidazoles, *Journal of Protozoology*, 39 (1992) 510–515.
- [10] B. Chavez, M. Espinosa-Cantellano, R. Cedillo-Rivera, A. Ramirez, A. Martinez-Palomo, Effects of albendazole on *Entamoeba histolytica* and *Giardia lamblia* trophozoites, *Archives of Medical Research*, 23 (1992) 63–67.
- [11] J. Valdez, R. Cedillo, A. Hernández-Campos, L. Yépez, F. Hernández-Luis, G. Navarrete-Vázquez, A. Tapia, R. Cortés, M. Hernández, R. Castillo, Synthesis and antiparasitic activity of 1*H*-benzimidazole derivatives, *Bioorganic & Medicinal Chemistry Letters*, 12 (2002) 2221–2224.
- [12] E. Lacey, Mode of action of benzimidazoles, *Parasitology Today*, 6 (1990) 112–115.
- [13] R. Sánchez Díaz, Tesis experimental, Maestría. Facultad de Química, UNAM, (1999).
- [14] A. Luna González, Tesis experimental, Maestría. Facultad de Química, UNAM, (2004).
- [15] G. Navarrete-Vázquez, R. Cedillo, A. Hernández-Campos, L. Yépez, F. Hernández-Luis, J. Valdez, R. Morales, R. Cortes, M. Hernández, R. Castillo, Synthesis and antiparasitic activity of 2-(trifluoromethyl)-benzimidazole derivatives, *Bioorganic & Medicinal Chemistry Letters*, 11 (2001) 187–190.
- [16] G. Navarrete-Vázquez, L. Yépez, A. Hernández-Campos, A. Tapia, F. Hernández-Luis, R. Cedillo, J. Gonzalez, A. Martínez-Fernandez, M. Martínez-Grueiro, R. Castillo, Synthesis and antiparasitic activity of albendazole and mebendazole analogues, *Bioorganic & Medicinal Chemistry*, 11 (2003) 4615–4622.

- [17] G. Navarrete-Vázquez, M.D. Rojano-Vilchis, L. Yépez-Mulia, V. Melendez, L. Gerena, A. Hernández-Campos, R. Castillo, F. Hernández-Luis, Synthesis and antiprotozoal activity of some 2-(trifluoromethyl)-1H-benzimidazole bioisosteres, *European Journal of Medicinal Chemistry*, 41 (2006) 135–141.
- [18] M. Andrzejewska, L. Yépez-Mulia, R. Cedillo-Rivera, A. Tapia, L. Vilpo, J. Vilpo, Z. Kazimierczuk, Synthesis, antiprotozoal and anticancer activity of substituted 2-trifluoromethyl- and 2-pentafluoroethylbenzimidazoles, *European Journal of Medicinal Chemistry*, 37 (2002) 973–978.
- [19] M. Andrzejewska, L. Yépez-Mulia, A. Tapia, R. Cedillo-Rivera, A.E. Laudy, B.J. Starosciak, Z. Kazimierczuk, Synthesis, and antiprotozoal and antibacterial activities of S-substituted 4,6-dibromo- and 4,6-dichloro-2-mercaptobenzimidazoles, *European Journal of Pharmaceutical Sciences*, 21 (2004) 323–329.
- [20] P. Vargas Benítez, Tesis experimental, Licenciatura, Universidad Autónoma Benito Juárez de Oaxaca.
- [21] D. Valdez-Padilla, S. Rodríguez-Morales, A. Hernández-Campos, F. Hernández-Luis, L. Yépez-Mulia, A. Tapia-Contreras, R. Castillo, Synthesis and antiprotozoal activity of novel 1-methylbenzimidazole derivatives, *Bioorganic & Medicinal Chemistry*, 17 (2009) 1724–1730.
- [22] F. Ooms, Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry, *Current Medicinal Chemistry*, 7 (2000) 141–158.
- [23] W.L. Jorgensen, The many roles of computation in drug discovery, *Science*, 303 (2004) 1813–1818.
- [24] F. López-Vallejo, J.L. Medina-Franco, R. Castillo, Diseño de fármacos asistido por computadora, *Educación Química* 17 (2006) 452–457.
- [25] W.L. Jorgensen, Efficient drug lead discovery and optimization, *Accounts of Chemical Research*, 42 (2009) 724–733.
- [26] A.J. Hopfinger, Computer-assisted drug design, *Journal of Medicinal Chemistry*, 28 (1985) 1133–1139.
- [27] G. Klebe, Recent developments in structure-based drug design, *Journal of Molecular Medicine-Jmm*, 78 (2000) 269–281.
- [28] D.B. Kitchen, H. Decornez, J.R. Furr, J. Bajorath, Docking and scoring in virtual screening for drug discovery: methods and applications, *Nature Reviews Drug Discovery*, 3 (2004) 935–949.
- [29] M.H.J. Seifert, K. Wolf, D. Vitt, Virtual high-throughput in silico screening, *BIOSILICO*, 1 (2003) 143–149.
- [30] H. Kubinyi, Comparative molecular field analysis (CoMFA), in: Wiley (Ed.) *Encyclopedia of computational Chemistry*, United Kingdom, 1998.
- [31] M. Sastry, J.F. Lowrie, S.L. Dixon, W. Sherman, Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments, *Journal of Chemical Information and Modeling*, 50 (2010) 771–784.
- [32] H. Kubinyi, QSAR and 3D QSAR in drug design .1. methodology, *Drug Discovery Today*, 2 (1997) 457–467.
- [33] H. Kubinyi, QSAR and 3D QSAR in drug design .2. Applications and problems, *Drug Discovery Today*, 2 (1997) 538–546.

- [34] J.L. Medina-Franco, S. Rodríguez-Morales, C. Juárez-Gordiano, A. Hernández-Campos, R. Castillo, Docking-based CoMFA and CoMSIA studies of non-nucleoside reverse transcriptase inhibitors of the pyridinone derivative type, *Journal of Computer-Aided Molecular Design*, 18 (2004) 345–360.
- [35] R.J. Spandl, A. Bender, D.R. Spring, Diversity-oriented synthesis; a spectrum of approaches and results, *Organic & Biomolecular Chemistry*, 6 (2008) 1149–1158.
- [36] R.A. Houghten, C. Pinilla, M.A. Giulianotti, J.R. Appel, C.T. Dooley, A. Nefzi, J.M. Ostresh, Y.P. Yu, G.M. Maggiora, J.L. Medina-Franco, D. Brunner, J. Schneider, Strategies for the use of mixture-based synthetic combinatorial libraries: Scaffold ranking, direct testing, in vivo, and enhanced deconvolution by computational methods, *Journal of Combinatorial Chemistry*, 10 (2008) 3–19.
- [37] B.R. Beno, J.S. Mason, The design of combinatorial libraries using properties and 3D pharmacophore fingerprints, *Drug Discovery Today*, 6 (2001) 251–258.
- [38] R.D. Brown, Y.C. Martin, Use of structure–activity data to compare structure-based clustering methods and descriptors for use in compound selection, *Journal of Chemical Information and Computer Sciences*, 36 (1996) 572–584.
- [39] V.J. Gillet, P. Willett, J. Bradshaw, The effectiveness of reactant pools for generating structurally-diverse combinatorial libraries, *Journal of Chemical Information and Computer Sciences*, 37 (1997) 731–740.
- [40] J.-L. Reymond, R. van Deursen, L.C. Blum, L. Ruddigkeit, Chemical space as a source for new drugs, *MedChemComm*, 1 (2010) 30–38.
- [41] R.D. Cramer, D.E. Patterson, J.D. Bunce, Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins, *Journal of the American Chemical Society*, 110 (1988) 5959–5967.
- [42] G. Klebe, U. Abraham, T. Mietzner, Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity, *Journal of Medicinal Chemistry*, 37 (1994) 4130–4146.
- [43] G. Klebe, U. Abraham, Comparative molecular similarity index analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries, *Journal of Computer-Aided Molecular Design*, 13 (1999) 1–10.
- [44] V.N. Viswanadhan, A.K. Ghose, G.R. Revankar, R.K. Robins, Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics, *Journal of Chemical Information and Computer Sciences*, 29 (1989) 163–172.
- [45] O. Temiz-Arpaci, B. Tekiner-Gulbas, I. Yildiz, E. Aki-Sener, I. Yalcin, QSAR-3D analysis on benzazole derivatives as eukaryotic topoisomerase II inhibitors by using comparative molecular field analysis method, *Bioorganic & Medicinal Chemistry*, 13 (2005) 6354–6359.
- [46] P.D. Patel, M.R. Patel, N. Kaushik-Basu, T.T. Talele, 3D QSAR and molecular docking studies of benzimidazole derivatives as hepatitis C virus NS5B polymerase inhibitors, *Journal of Chemical Information and Modeling*, 48 (2008) 42–55.
- [47] M.L. López-Rodríguez, M. Murcia, B. Benhamu, A. Viso, M. Campillo, L. Pardo, Benzimidazole derivatives. 3. QSAR-3D/CoMFA model and computational

- simulation for the recognition of 5-HT₄ receptor antagonists, *Journal of Medicinal Chemistry*, 45 (2002) 4806–4815.
- [48] F. López-Vallejo, J.L. Medina-Franco, A. Hernández-Campos, S. Rodríguez-Morales, L. Yépez, R. Cedillo, R. Castillo, Molecular modeling of some 1*H*-benzimidazole derivatives with biological activity against *Entamoeba histolytica*: A comparative molecular field analysis study, *Bioorganic & Medicinal Chemistry*, 15 (2007) 1117–1126.
- [49] A. Golbraikh, A. Tropsha, Beware of q(2)!, *Journal of Molecular Graphics & Modelling*, 20 (2002) 269–276.
- [50] A. Golbraikh, A. Tropsha, Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection, *Journal of Computer-Aided Molecular Design*, 16 (2002) 357–369.
- [51] S. Rodríguez-Morales, Informe de actividades de estancia posdoctoral, Facultad de Química, UNAM.
- [52] G.M. Maggiora, On outliers and activity cliffs-Why QSAR often disappoints, *Journal of Chemical Information Modeling*, 46 (2006) 1535–1535.
- [53] J. Bajorath, L. Peltason, M. Wawer, R. Guha, M.S. Lajiness, J.H. Van Drie, Navigating structure–activity landscapes, *Drug Discovery Today*, 14 (2009) 698–705.
- [54] R. Guha, J.H. Van Drie, Structure–activity landscape index: identifying and quantifying *activity cliffs*, *Journal of Chemical Information and Modeling*, 48 (2008) 646–658.
- [55] J.L. Medina-Franco, K. Martinez-Mayorga, A. Bender, R.M. Marin, M.A. Giulianotti, C. Pinilla, R.A. Houghten, Characterization of activity landscapes using 2D and 3D similarity methods: consensus activity cliffs, *Journal of Chemical Information and Modeling*, 49 (2009) 477–491.
- [56] Spartan, Version 02, Wavefunction Inc., Irvine, CA.
- [57] J.J.P. Stewart, Optimization of parameters for semiempirical methods .1. method, *Journal of Computational Chemistry*, 10 (1989) 209–220.
- [58] Molecular Operating Environment (MOE), Version 2007; Chemical Computing Group Inc.: Montreal, Quebec, Canada.
- [59] Canvas 1.3, Schrodinger, LLC, New York, NY 2010.
- [60] P. Jaccard, *Société Vaudoise des Sciences Naturelles*, 37 (1901) 547–579.
- [61] P. Willett, J.M. Barnard, G.M. Downs, Chemical similarity searching, *Journal of Chemical Information and Computer Sciences*, 38 (1998) 983–996.
- [62] L. Peltason, J. Bajorath, Molecular similarity analysis uncovers heterogeneous structure-activity relationships and variable activity landscapes, *Chemistry and Biology*, 14 (2007) 489–497.
- [63] R.S. Mulliken, Electronic population analysis on LCAO-MO molecular wave functions, *Journal of Chemical Physics*, 23 (1955) 1833–1840.
- [64] S.R. Cox, D.E. Williams, Representation of the molecular electrostatic potential by a net atomic charge model, *Journal of Computational Chemistry*, 2 (1981) 304–323.
- [65] J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity - a rapid access to atomic charges, *Tetrahedron*, 36 (1980) 3219–3228.
- [66] Sybyl, Tripos Inc., Version 7.3, St. Louis, MO.
- [67] ROCS, OpenEye Scientific Software Inc., Version 2.3.1, Santa Fe, NM.

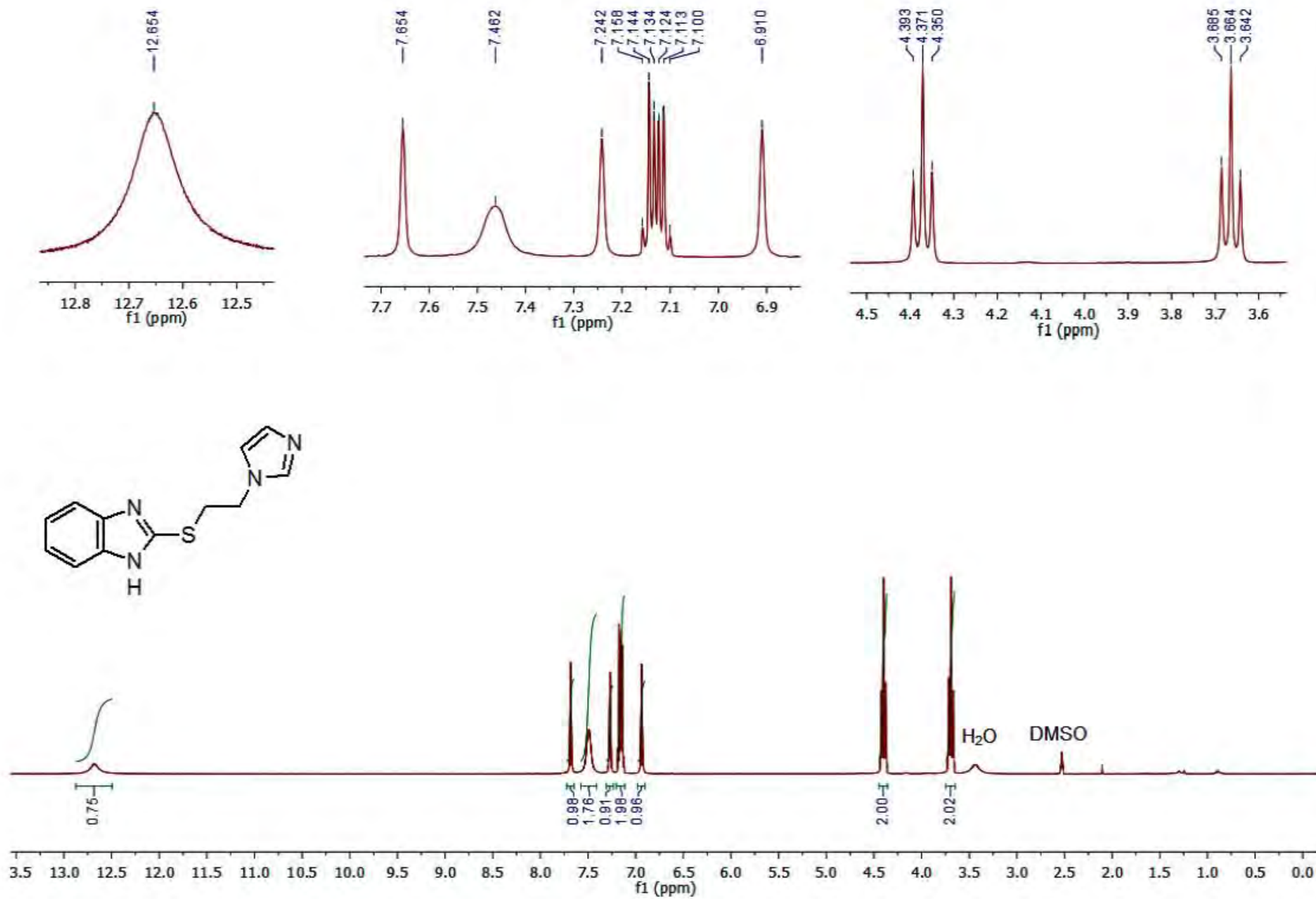
- [68] T.S. Rush, J.A. Grant, L. Mosyak, A. Nicholls, A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction, *Journal of Medicinal Chemistry*, 48 (2005) 1489–1495.
- [69] M.J. Sykes, M.J. Sorich, J.O. Miners, Molecular Modeling Approaches for the Prediction of the Nonspecific Binding of Drugs to Hepatic Microsomes, *Journal Chemical Information and Modeling*, 46 (2006) 2661–2673.
- [70] J.A.D. VanAllan, B. D., 2-Mercaptobenzimidazole, *Organic Syntheses, Coll. Vol.*, 4 569.
- [71] D. Bogdal, K. Jaskot, Synthesis of vinyl monomers with active azaaromatic groups. Phase-transfer catalytic approach, *Synthetic Communications*, 30 (2000) 3341–3352.
- [72] T. Scior, J.L. Medina-Franco, Q.T. Do, K. Martinez-Mayorga, J.A. Yunes Rojas, P. Bernard, How to recognize and workaroud pitfalls in QSAR studies: a critical review, *Current Medicinal Chemistry*, 16 (2009) 4297–4313.
- [73] A. Bender, J.L. Jenkins, J. Scheiber, S.C.K. Sukuru, M. Glick, J.W. Davies, How similar are similarity searching methods? a principal component analysis of molecular descriptor space, *Journal of Chemical Information and Modeling*, 49 (2009) 108–119.
- [74] A. Bender, How similar are those molecules after all? Use two descriptors and you will have three different answers, *Expert Opinion on Drug Discovery*, 5 (2010) 1141–1151.
- [75] A.B. Yongye, K. Byler, R. Santos, K. Martinez-Mayorga, G.M. Maggiora, J.L. Medina-Franco, Consensus models of activity landscapes with multiple chemical, conformer, and property representations, *Journal of Chemical Information and Modeling*, 51 (2011) 1259–1270.
- [76] D. Rogers, M. Hahn, Extended-connectivity fingerprints, *Journal of Chemical Information and Modeling*, 50 (2010) 742–754.
- [77] L. Peltason, P. Iyer, J.r. Bajorath, Rationalizing three-dimensional activity landscapes and the influence of molecular representations on landscape topology and the formation of activity cliffs, *Journal of Chemical Information and Modeling*, 50 (2010) 1021–1033.
- [78] H. Eckert, J. Bajorath, Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches, *Drug Discovery Today*, 12 (2007) 225–233.
- [79] P. Willett, Similarity-based virtual screening using 2D fingerprints, *Drug Discovery Today*, 11 (2006) 1046–1053.
- [80] S.S. Kulkarni, A.H. Newman, W.J. Houlihan, Three-dimensional quantitative structure-activity relationships of mazindol analogues at the dopamine transporter, *Journal of Medicinal Chemistry*, 45 (2002) 4119–4127.
- [81] V. Ravichandran, R.K. Agrawal, Predicting anti-HIV activity of PETT derivatives: CoMFA approach, *Bioorganic & Medicinal Chemistry Letters*, 17 (2007) 2197–2202.
- [82] A.K. Chakraborti, B. Gopalakrishnan, M.E. Sobhia, A. Malde, Comparative molecular field analysis (CoMFA) of phthalazine derivatives as phosphodiesterase IV inhibitors, *Bioorganic & Medicinal Chemistry Letters*, 13 (2003) 2473–2479.
- [83] Y.C. Martin, Let's not forget tautomers, *Journal of Computer-Aided Molecular Design*, 23 (2009) 693–704.

- [84] A.R. Katritzky, C.D. Hall, B. El-Dien, M. El-Gendy, B. Draghici, Tautomerism in drug discovery, *Journal of Computer-Aided Molecular Design*, 24 (2010) 475–484.
- [85] F. Milletti, A. Vulpetti, Tautomer preference in PDB complexes and its impact on structure-based drug discovery, *Journal of Chemical Information and Modeling*, 50 (2010) 1062–1074.
- [86] K.H. Kim, Outliers in SAR and QSAR: Is unusual binding mode a possible source of outliers?, *Journal of Computer-Aided Molecular Design*, 21 (2007) 63–86.
- [87] K.H. Kim, Outliers in SAR and QSAR: 2. Is a flexible binding site a possible source of outliers?, *Journal of Computer-Aided Molecular Design*, 21 (2007) 421–435.
- [88] M.T. Sisay, L. Peltason, J. Bajorath, Structural interpretation of activity cliffs revealed by systematic analysis of structure-activity relationships in analog series, *Journal Chemical Information and Modeling*, 49 (2009) 2179–2189.
- [89] M. Boiani, M. Gonzalez, Imidazole and benzimidazole derivatives as chemotherapeutic agents, *Mini-Reviews in Medicinal Chemistry*, 5 (2005) 409–424.
- [90] J. Horn, The proton-pump inhibitors: similarities and differences, *Clinical Therapeutics*, 22 (2000) 266–280.
- [91] C.R. Chong, D.J. Sullivan, New uses for old drugs, *Nature*, 448 (2007) 645–646.
- [92] T.T. Ashburn, K.B. Thor, Drug repositioning: identifying and developing new uses for existing drugs, *Nature Reviews Drug Discovery*, 3 (2004) 673–683.
- [93] K.A. O'Connor, B.L. Roth, Finding new tricks for old drugs: an efficient route for public-sector drug discovery, *Nature Reviews Drug Discovery*, 4 (2005) 1005–1014.
- [94] E. Dolgin, Nonprofit disease groups earmark grants for drug repositioning, *Nature Medicine*, 17 (2011) 1027–1027.
- [95] R. Galbreath, R. Ingham, Synthesis of 2-chlorophenonthiazine via a smiles rearrangement, *Journal of Organic Chemistry*, 23 (1958) 1804–1806.
- [96] J.T. Manka, F. Guo, J. Huang, H. Yin, J.M. Farrar, M. Sienkowska, V. Benin, P. Kaszynski, Synthesis of polyfunctionalized biphenyls as intermediates for a new class of liquid crystals, *Journal of Organic Chemistry*, 68 (2003) 9574–9588.
- [97] I. Yildiz-Oren, I. Yalcin, E. Aki-Sener, N. Ucarturk, Synthesis and structure–activity relationships of new antimicrobial active multisubstituted benzazole derivatives, *European Journal of Medicinal Chemistry*, 39 (2004) 291–298.
- [98] E.M. Fry, 2-Hydroxycyclohexaneacetic acid lactones as intermediates in the synthesis of hydrogenated phenanthrenes, *Journal of Organic Chemistry*, 17 (1952) 1484–1491.
- [99] H.D. Thakuria, G., An expeditious one-pot solvent free synthesis of benzimidazole derivatives, *ARKIVOC*, (2008) 321–328.
- [100] S. Rodríguez-Morales, R. Sánchez-Díaz, A. Hernández-Martínez, A. Tapia, R. Cedillo, M. Hernández, A. Hernández-Campos, F. Hernández-Luis, L. Yépez-Mulia, R. Castillo, *Manuscrito en preparación*
- [101] L. Spirer, Preparation of benzimidazole derivatives, *Roczniki Chemii*, 28 (1954) 455.
- [102] D. Kalyani, A.R. Dick, W.Q. Anani, M.S. Sanford, Scope and selectivity in palladium-catalyzed directed C–H bond halogenation reactions, *Tetrahedron*, 62 (2006) 11483–11498.

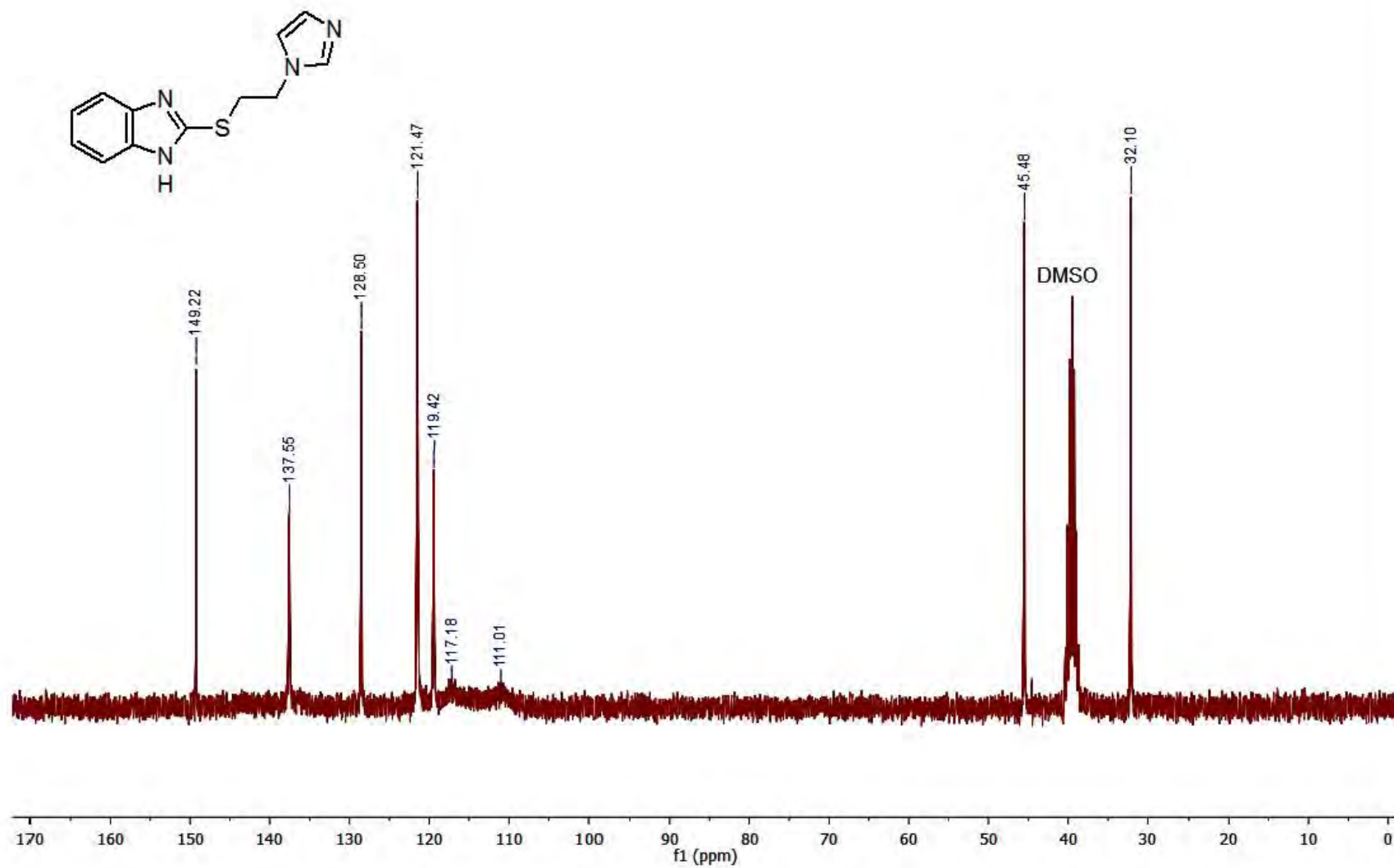
- [103] F.H. Case, S. Catino, F. Scholnick, Substituted 1,10-phenantrolines. VI. Chloro derivatives¹, *Journal of Organic Chemistry*, 19 (1954) 31–36.
- [104] D.W. Woolley, A. Pringle, Relationship of chemical structure to antibacterial activity among analogues of dimethyldiaminobenzene, *Journal of Biological Chemistry*, 194 (1952) 729–746.
- [105] N.V. Zotova, P.M. Kushakova, V.A. Kuznetsov, A.A. Rodin, A.V. Garabadzhiu, Specific features of nucleophilic substitution in 1-chloro-3,4-dinitrobenzene, *Russian Journal of Organic Chemistry*, 40 (2004) 1473–1476.
- [106] J. Rosevear, J. Wilshire, A comparison of the reactions of some ethyl *N*-arylcaramates with those of the corresponding acetanilides. I. Nitration, *Australian Journal of Chemistry*, 38 (1985) 723–733.
- [107] D. Schelz, N. Rotzler, Über die synthese von 1-aryl- und 1-alkyl-2,3-dimethyl-chinoxalinium-perchloraten: Teil 3-synthese von 1,2,3-trimethyl-6-X- und 1-cyclohexyl-2,3-dimethyl-6-X-chinoxalinium-perchloraten, *Dyes and Pigments*, 4 (1983) 305–320.
- [108] M.A. Xiang, P.J. Rybczynski, M. Patel, R.H. Chen, D.F. McComsey, H.-C. Zhang, J.W. Gunnet, R. Look, Y. Wang, L.K. Minor, H.M. Zhong, F.J. Villani, K.T. Demarest, B.P. Damiano, B.E. Maryanoff, Next-generation spirobenzazepines: Identification of RWJ-676070 as a balanced vasopressin V1a/V2 receptor antagonist for human clinical studies, *Bioorganic & Medicinal Chemistry Letters*, 17 (2007) 6623–6628.
- [109] I.A. Ismail, D.E. Sharp, M.R. Chedekel, Synthesis of benzothiazoles. α -Amino-(4-hydroxy-6-benzothiazolyl)propionic acid, *Journal of Organic Chemistry*, 45 (1980) 2243–2246.
- [110] D. Sperandio, A.R. Gangloff, J. Litvak, R. Goldsmith, J.M. Hataye, V.R. Wang, E.J. Shelton, K. Elrod, J.W. Janc, J.M. Clark, K. Rice, S. Weinheimer, K.-S. Yeung, N.A. Meanwell, D. Hernandez, A.J. Staab, B.L. Venables, J.R. Spencer, Highly potent non-peptidic inhibitors of the HCV NS3/NS4A serine protease, *Bioorganic and Medicinal Chemistry Letters*, 12 (2002) 3129–3133.
- [111] K. Nagarajan, A. Venkateswarlu, C.L. Kulkarni, A.N. Goud, R.K. Shah, Condensed heterotricycles. Amino and aminoalkyldibenz[b,f][1,4]oxazepin-11(10*H*)-ones, *Indian Journal of Chemistry* 12 (1974) 236–246.
- [112] J. Srogl, J. Hyvl, A. Revesz, D. Schroder, Mechanistic insights into a copper-disulfide interaction in oxidation of imines by disulfides, *Chemical Communications* (2009) 3463–3465.
- [113] Acetals, in, E. I. du Pont de Nemours & Co., GB695164, 1953.
- [114] Y. Migita, T. Eguchi, Y. Kumazawa, J. Nakagami, T. Amano, K. Sota, J. Sakakibara, Benzo[a]phenazine derivatives, with antitumor activity, and a process for their preparation, in, Taisho Pharmaceutical Co., Ltd., Japan . 1986, 30.
- [115] J. Miller, The SN mechanism in aromatic compounds. VI. Carbonyl and nitrile substituents, *Journal of the American Chemical Society*, 76 (1954) 448–450.
- [116] S. Oezden, D. Atabey, S. Yildiz, H. Goeker, Synthesis and potent antimicrobial activity of some novel methyl or ethyl 1*H*-benzimidazole-5-carboxylates derivatives carrying amide or amidine groups, *Bioorganic and Medicinal Chemistry*, 13 (2005) 1587–1597.
- [117] F. Reverdin, Nitration of the methyl ester of nitro-3-*p*-dimethylaminobenzoic acid, *Bulletin de la Societe Chimique de France*, 3 (1908) 126–133.

- [118] S.S. Gandhi, M.S. Gibson, M.L. Kaldas, S.M. Vines, Reactions of some aromatic nitro compounds with alkali metal amides, *Journal of Organic Chemistry*, 44 (1979) 4705–4707.
- [119] S. Sadaphal, V. Markhele, S. Sonar, M. Shingare, Silica supported sulphamic acid as mild catalyst for synthesis of anilides (Beckmann Rearrangement), using microwave irradiation, *Journal of the Korean Chemical Society*, 52 (2008) 454–456.
- [120] H. Oelschlaeger, Analysis of drugs by polarography and oscillopolarography. I. Polarographic determination of phenacetin, *Archiv der Pharmazie und Berichte der Deutschen Pharmazeutischen Gesellschaft*, 296 (1963) 7–13.
- [121] H. Oelsehlaeger, C. Welsch, Detection of phenacetin by nitration, *Archiv der Pharmazie und Berichte der Deutschen Pharmazeutischen Gesellschaft*, 297 (1964) 608–617.
- [122] H.F. Wilson, D.H. McRae, B.M. Vittimberga, Substituted *N*-aryllactams, in, Rohm & Haas Co. 1966, 11.
- [123] H.E. Ungnade, I. Ortega, Nitration and bromination of *m*-phenoxyanisole, *Journal Organic Chemistry*, 17 (1952) 1475–1483.
- [124] K.B. Mathur, A.P. Bhaduri, R.N. Iyer, N.M. Khanna, M.L. Dhar, Potential antimycobacterial agents. XVII. Synthesis of some cyclic analogs of thiocarbanilides, *Indian Journal of Chemistry*, 3 (1965) 397–401.
- [125] A.N. Lomakin, A.M. Simonov, V.A. Chigrina, Benzimidazole derivatives. XII. Action of sodium amide on *N*-alkylbenzimidazoles substituted in the 5-position, *Zhurnal Obshchei Khimii*, 33 (1963) 204–207.
- [126] K.J.P. Orton, H. King, Chlorination of anilines and phenols, *Journal of the Chemical Society, Perkin Transactions*, 99 (1911) 1185–1192.
- [127] E.H. Van, Halogenated nitrophenols, *Journal für Praktische Chemie (Leipzig)*, 127 (1930) 20–38.

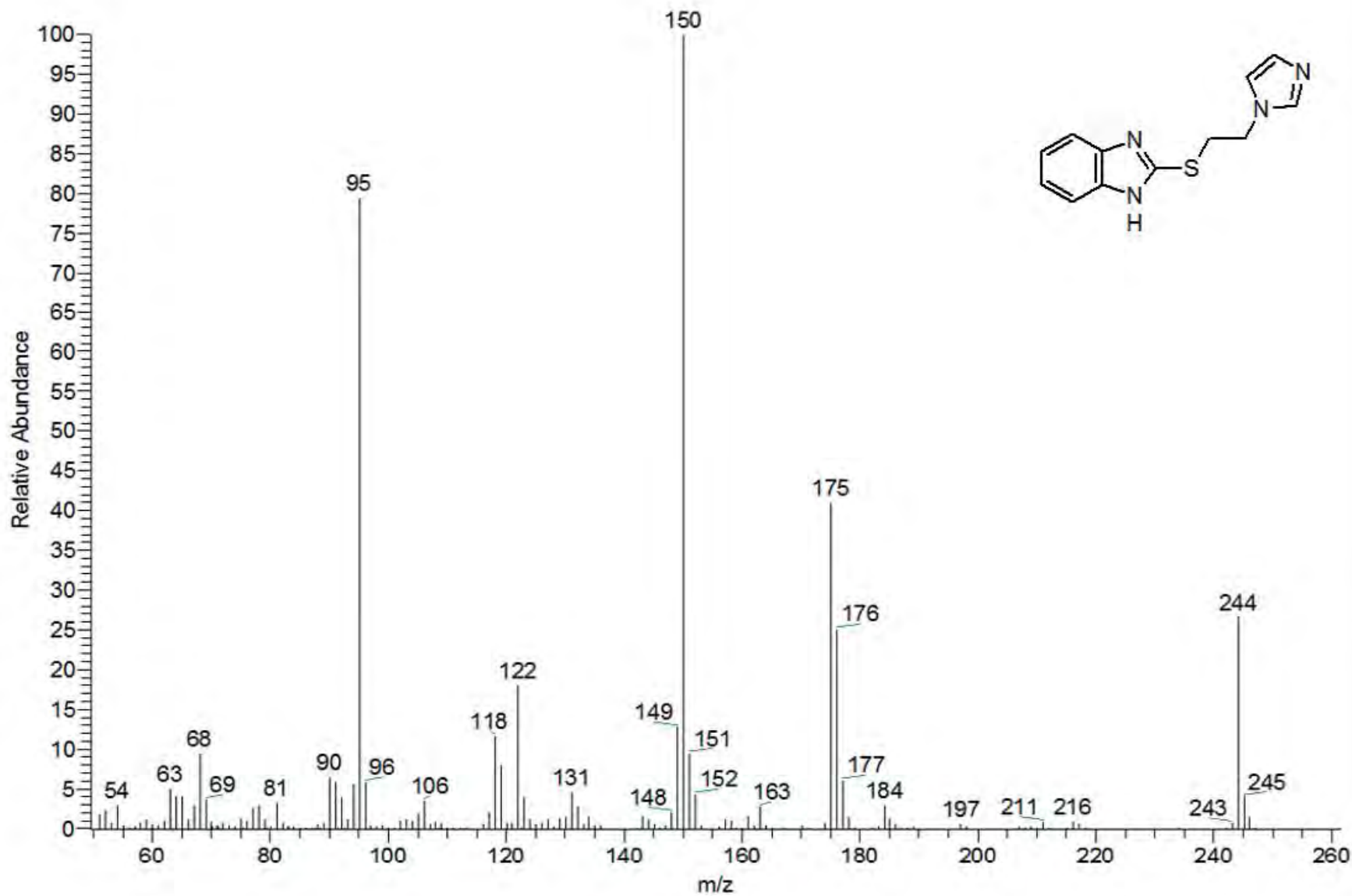
ESPECTROS



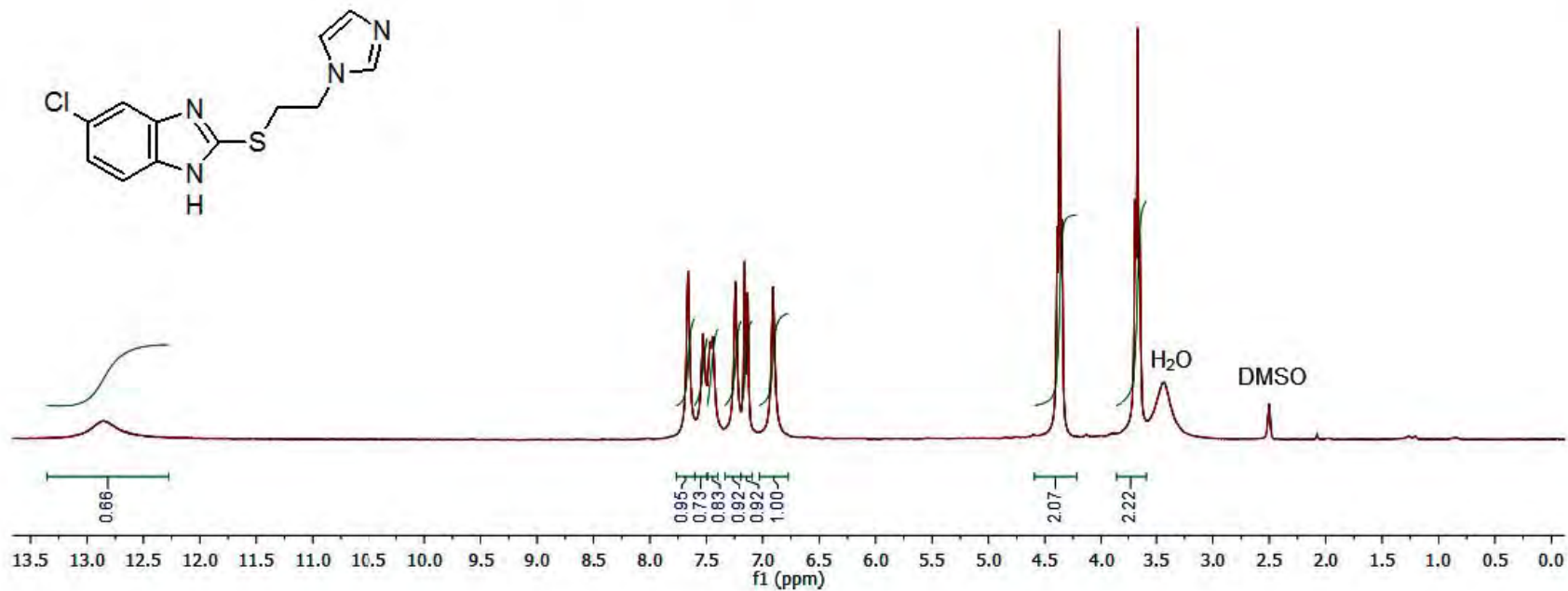
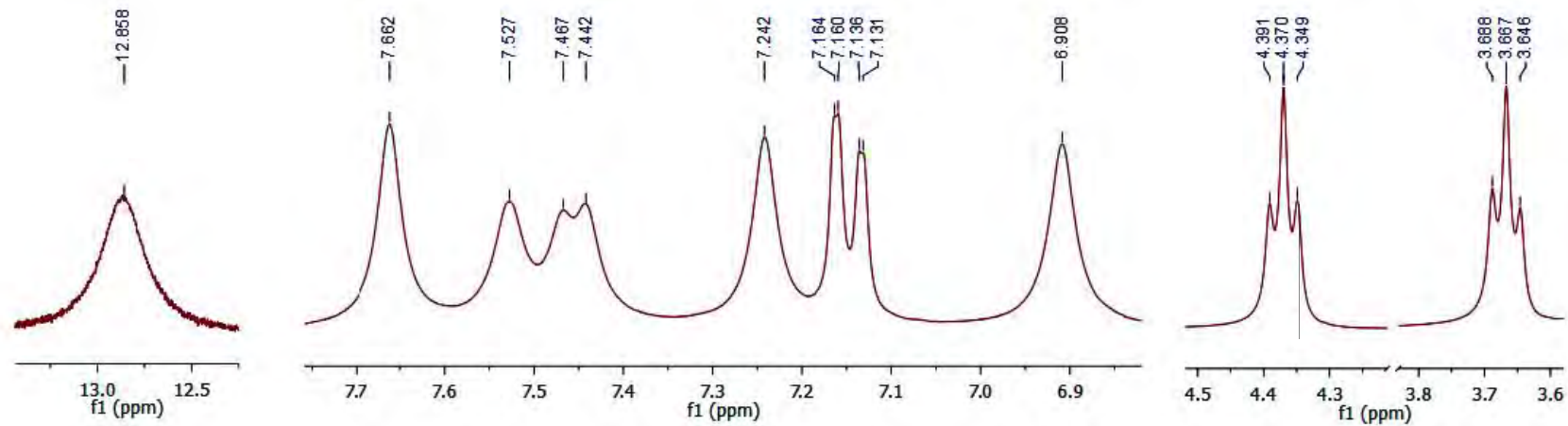
Espectro 1. RMN ^1H (300 MHz) del 2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1H-bencimidazol (S1J01)



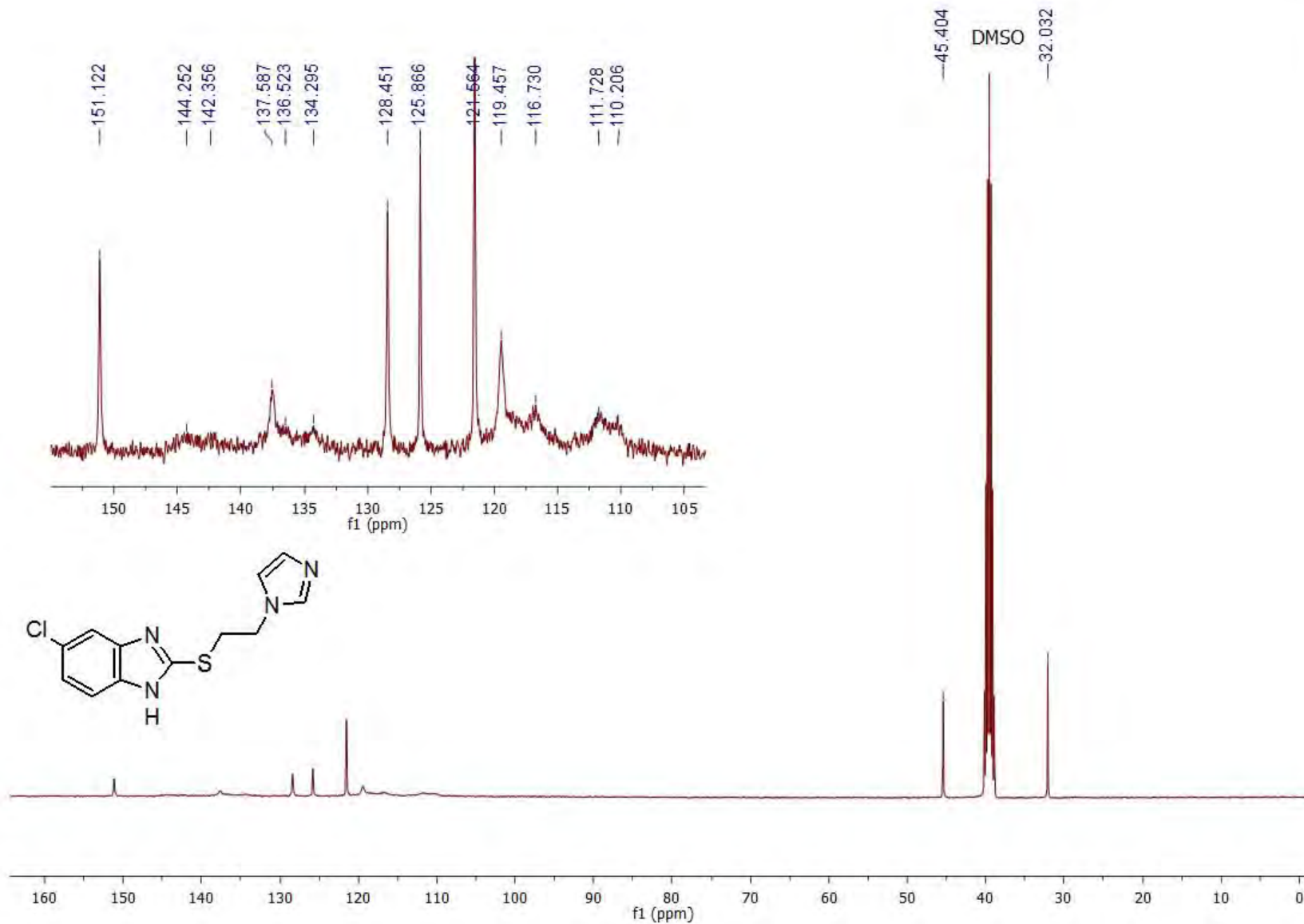
Espectro 2. RMN ¹³C (75.5 MHz) del 2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (**S1J01**)

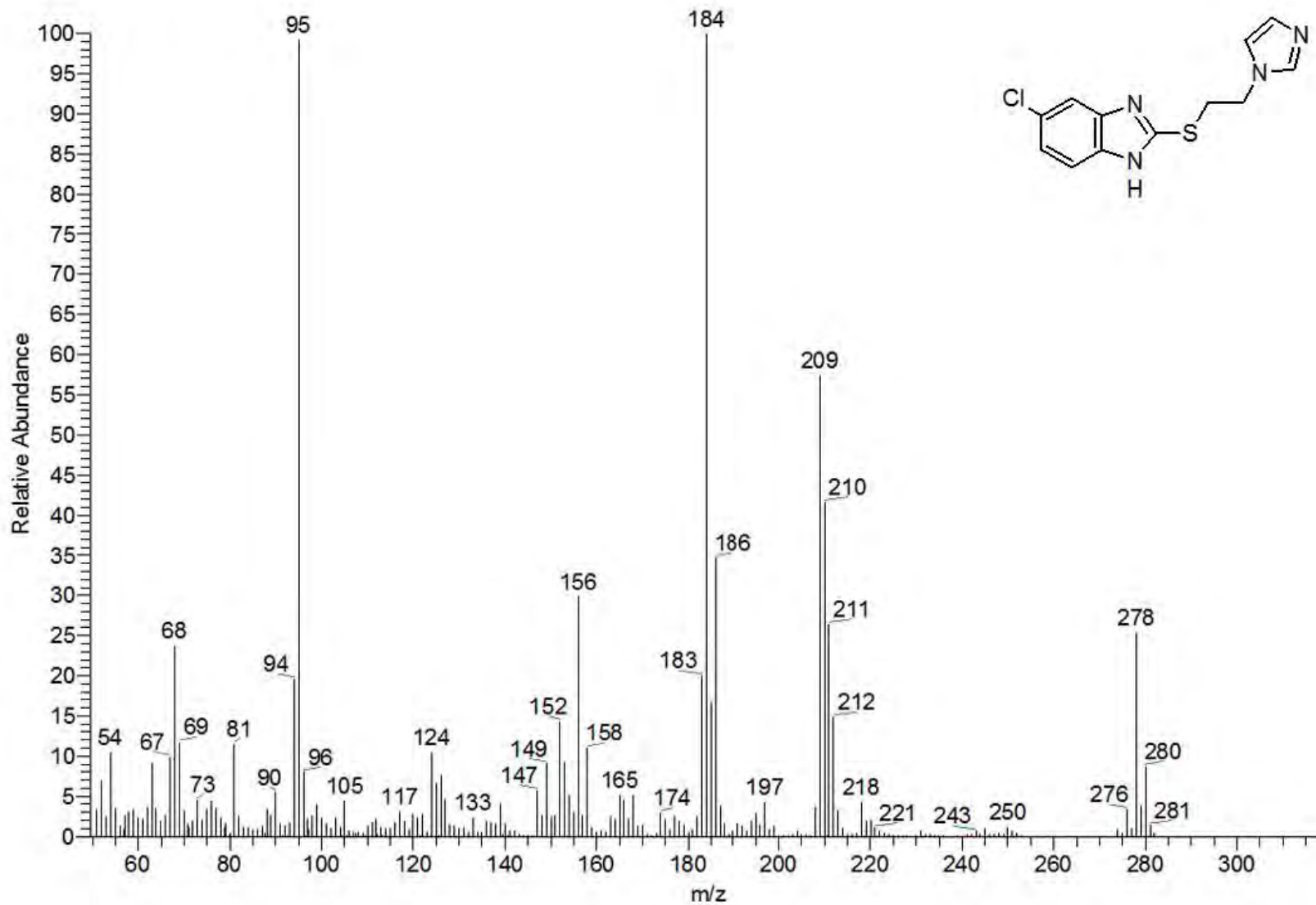


Espectro 3. EM (IE) 2-([2-(1H-imidazol-1-yl)etil]sulfanil)-1H-bencimidazol (S1J01)

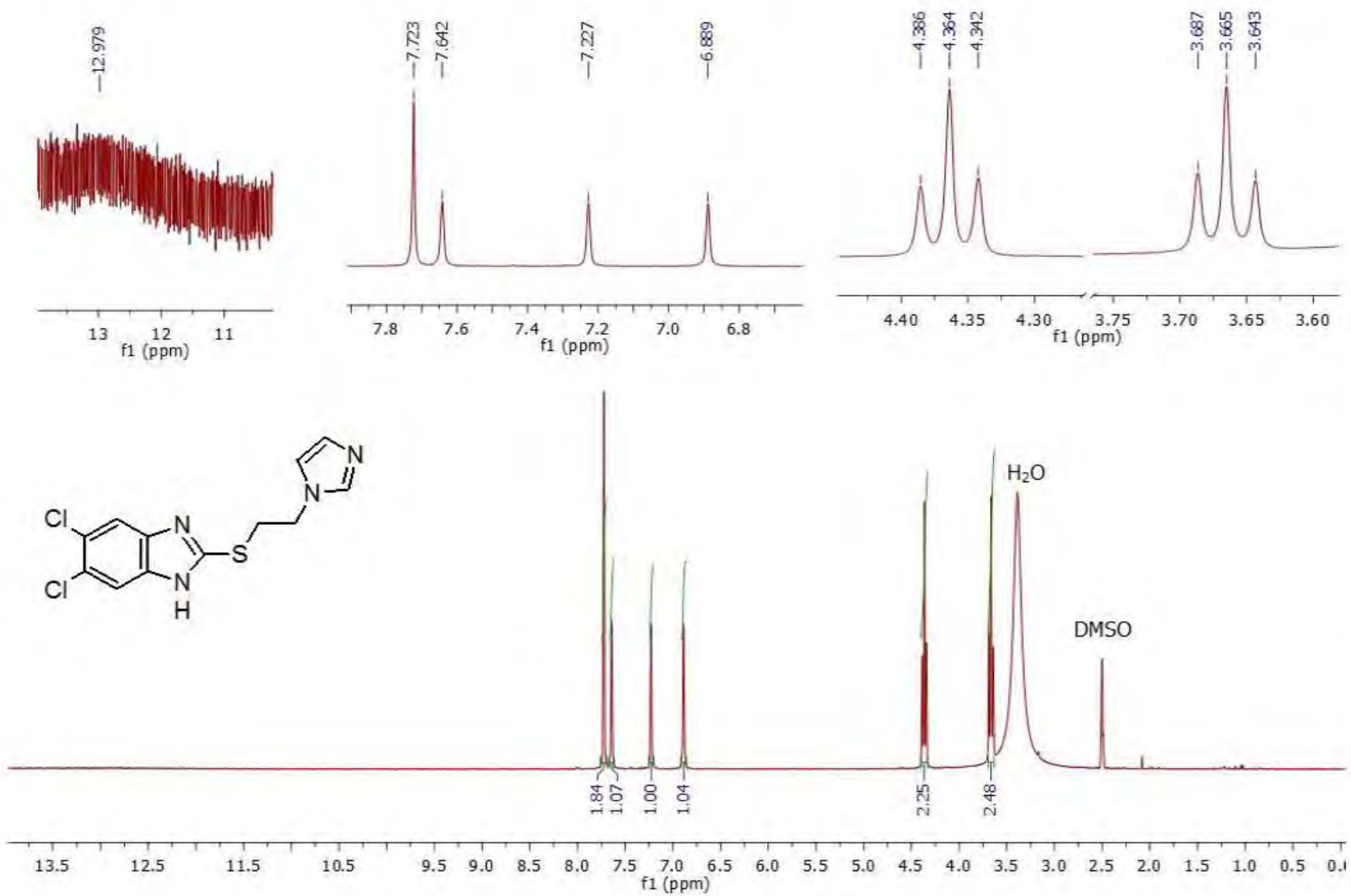


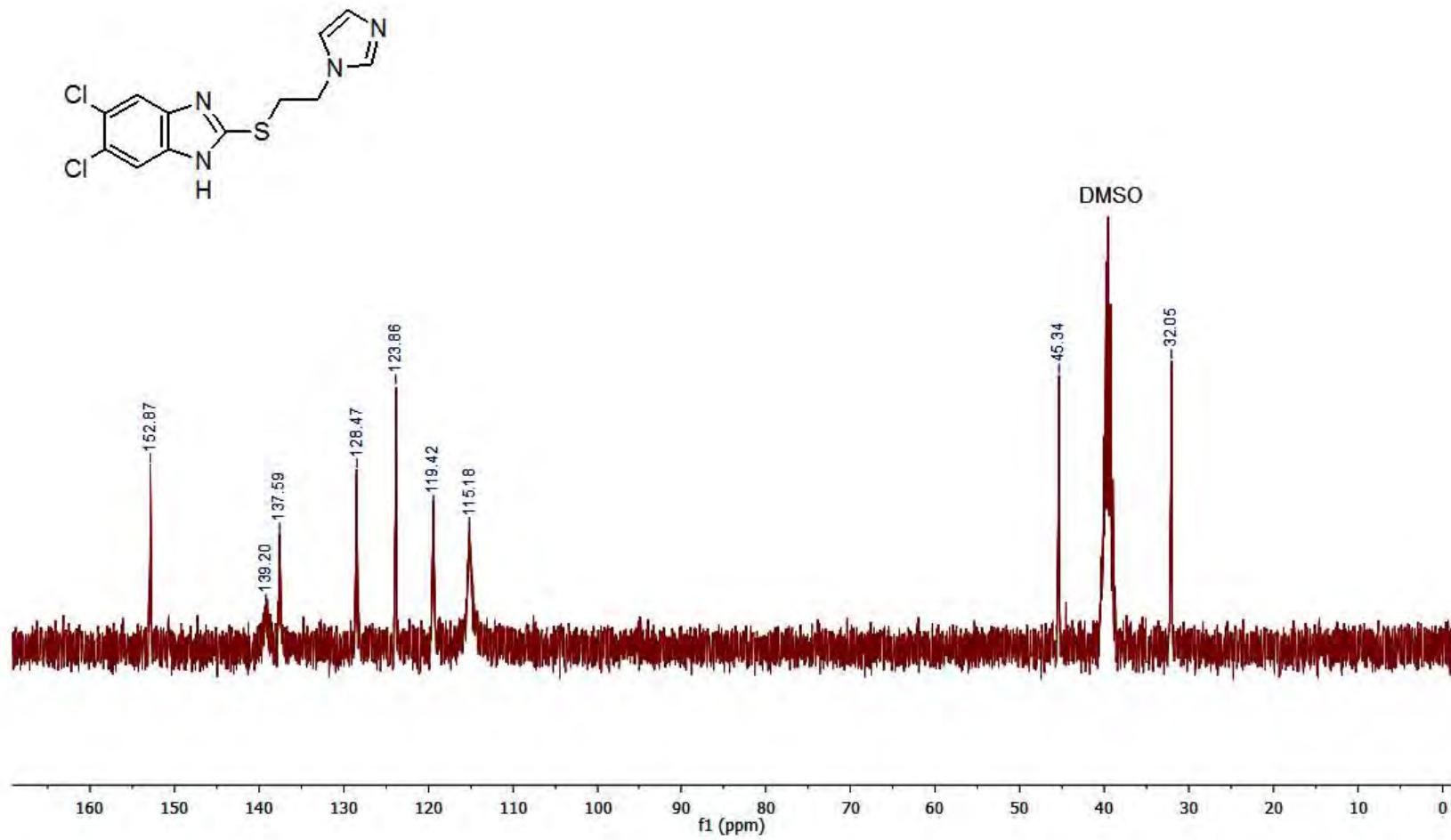
Espectro 4. RMN ^1H (300 MHz) del 5-cloro-2-([2-(1H-imidazol-1-il)etil]sulfanil)-1H-bencimidazol (S1J02)

Espectro 5. RMN ^{13}C (75.5 MHz) del 5-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (**S1J02**)

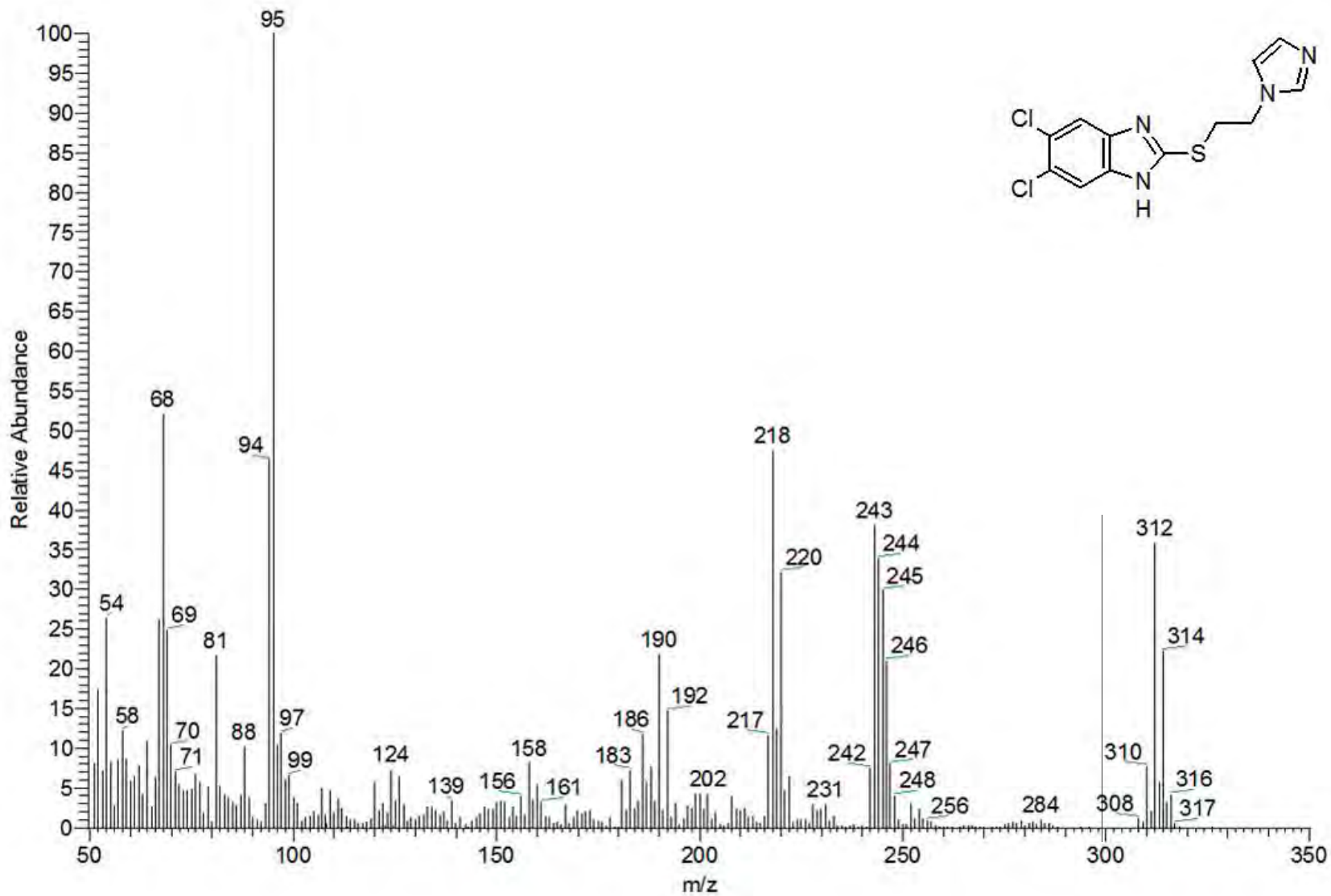


Espectro 6. EM (IE) 5-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (S1J02)

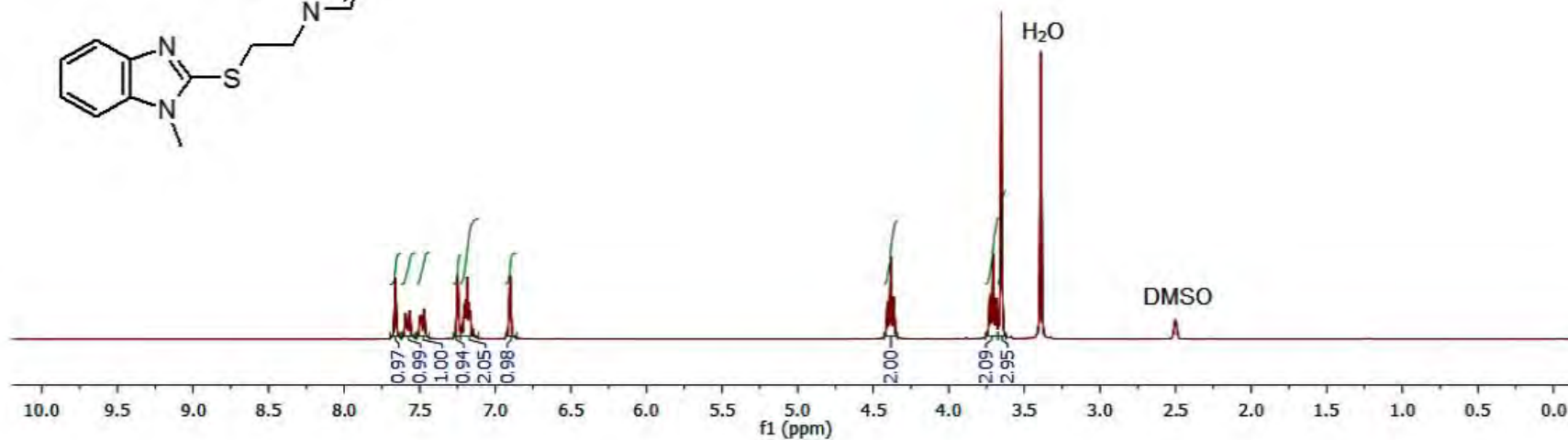
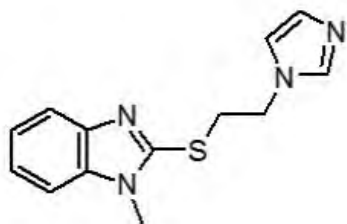
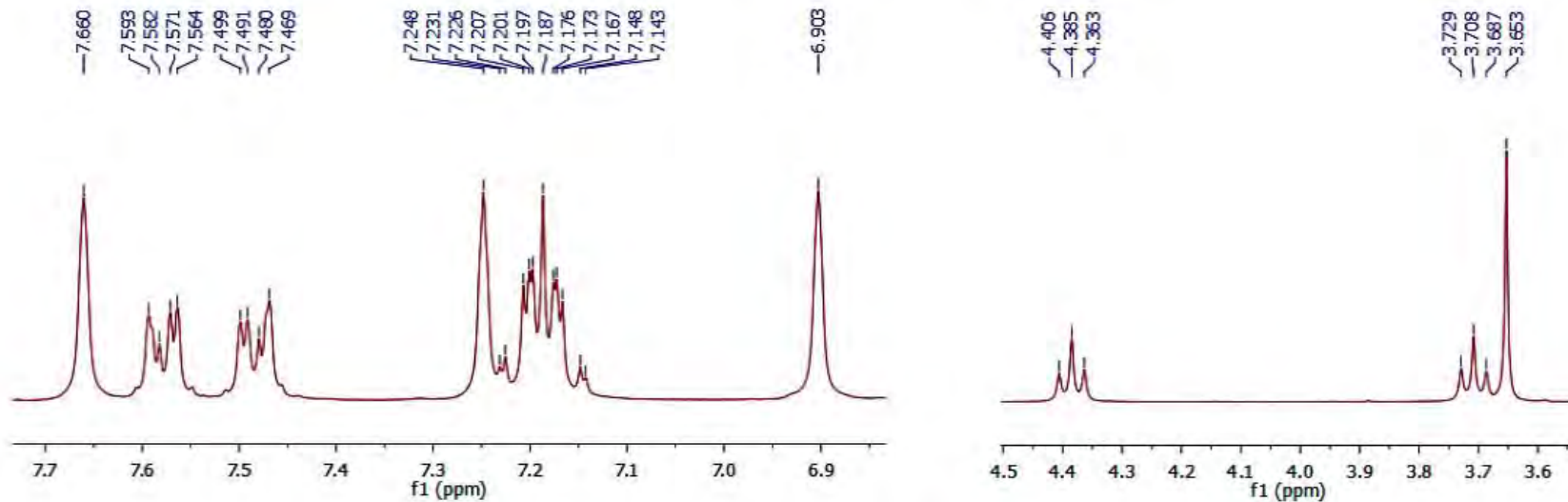
Espectro 7. RMN ¹H (300 MHz) del 5,6-dicloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-benzimidazol (**S1J03**)



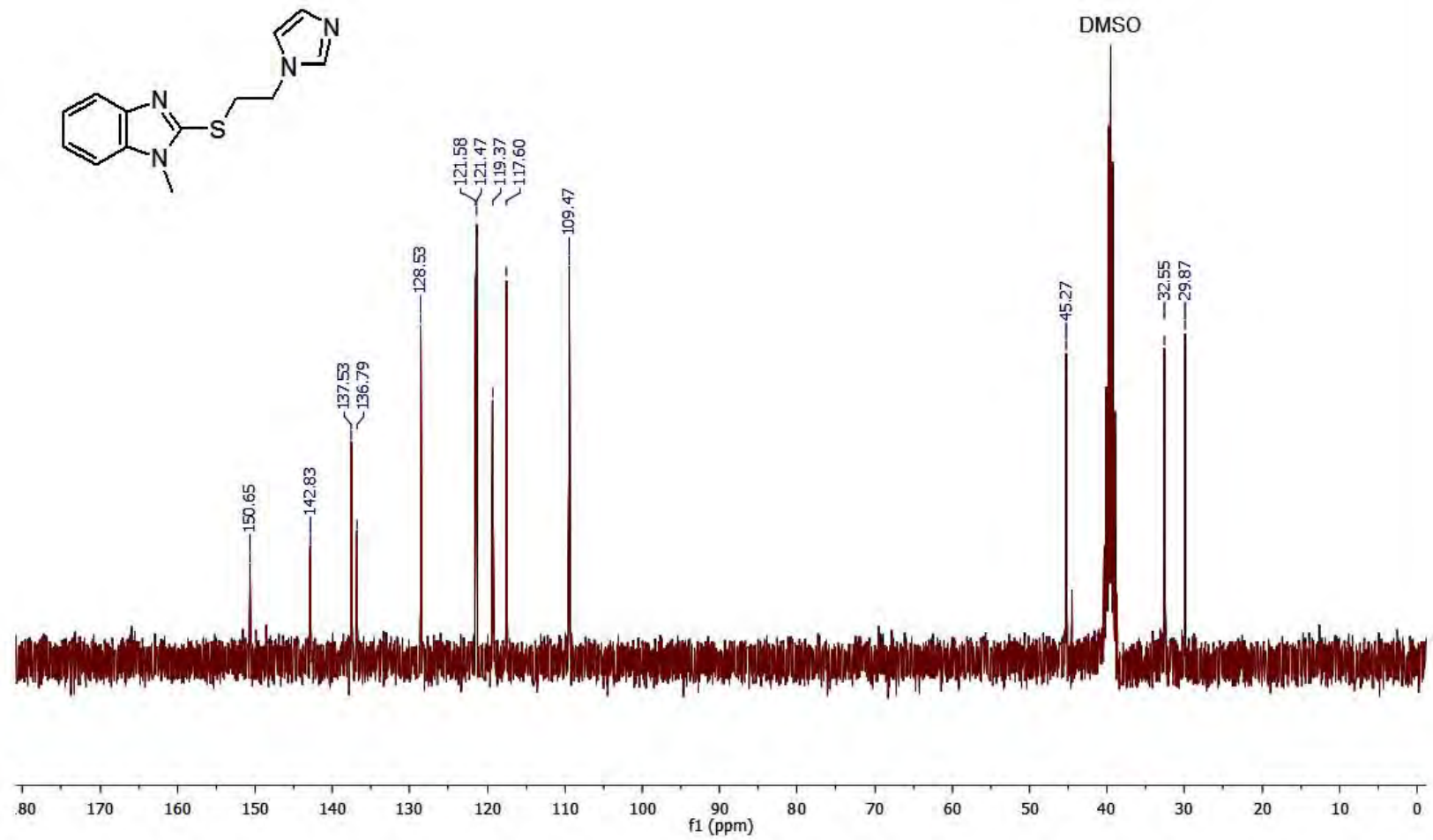
Espectro 8. RMN ^{13}C (75.5 MHz) del 5,6-dicloro-2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1H-benzimidazol (S1J03)

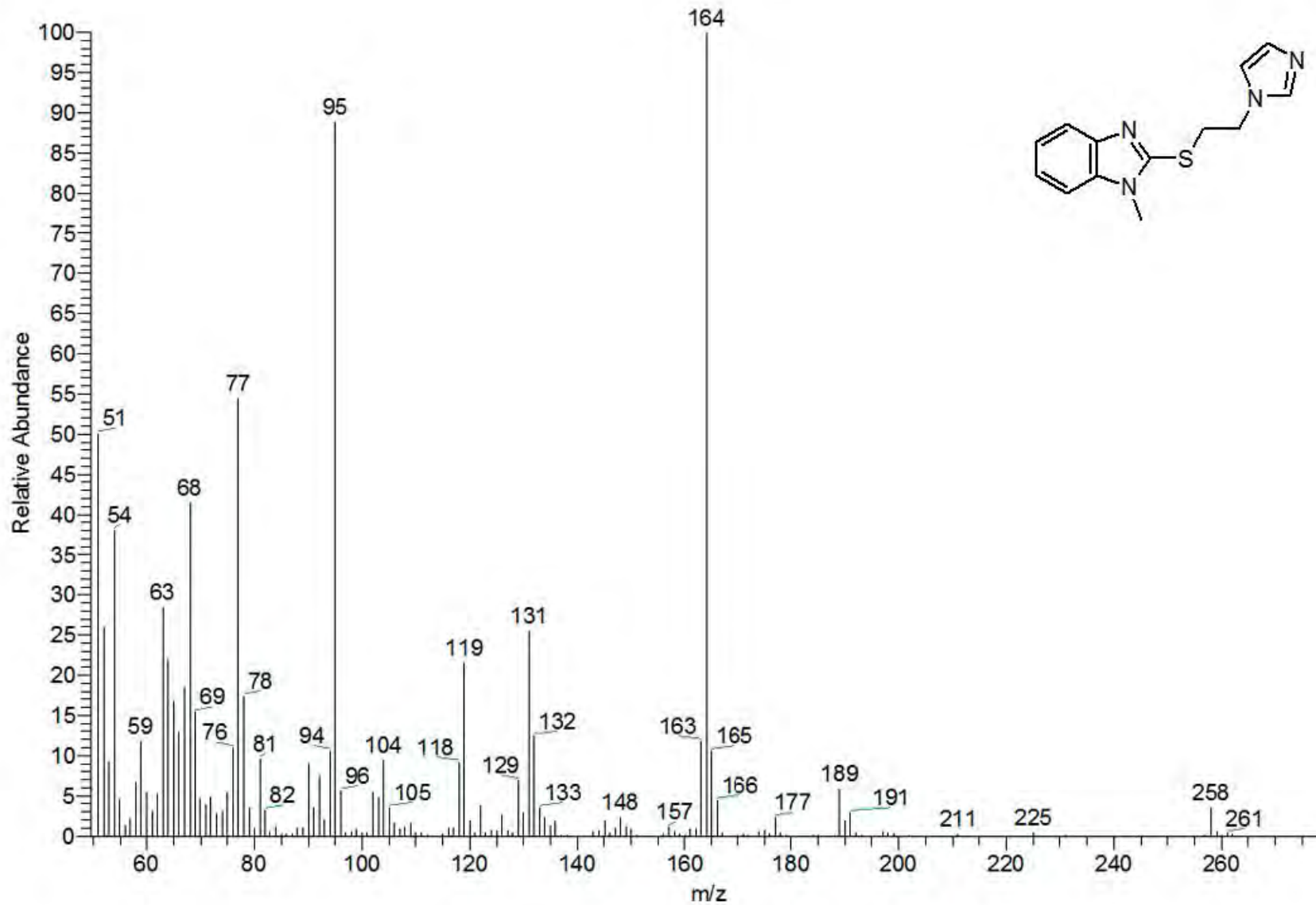


Espectro 9. EM (IE) 5,6-dicloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (S1J03)

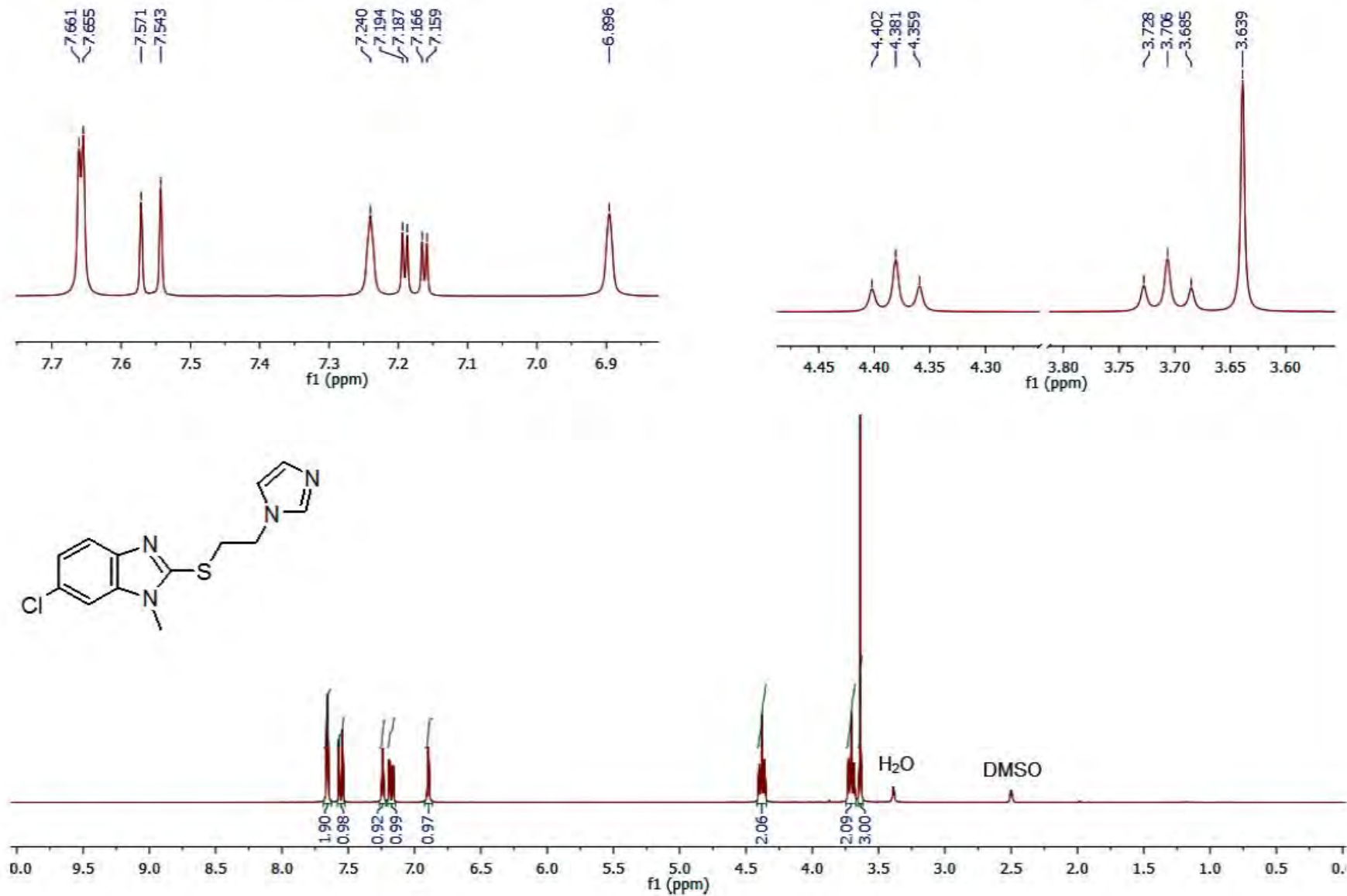


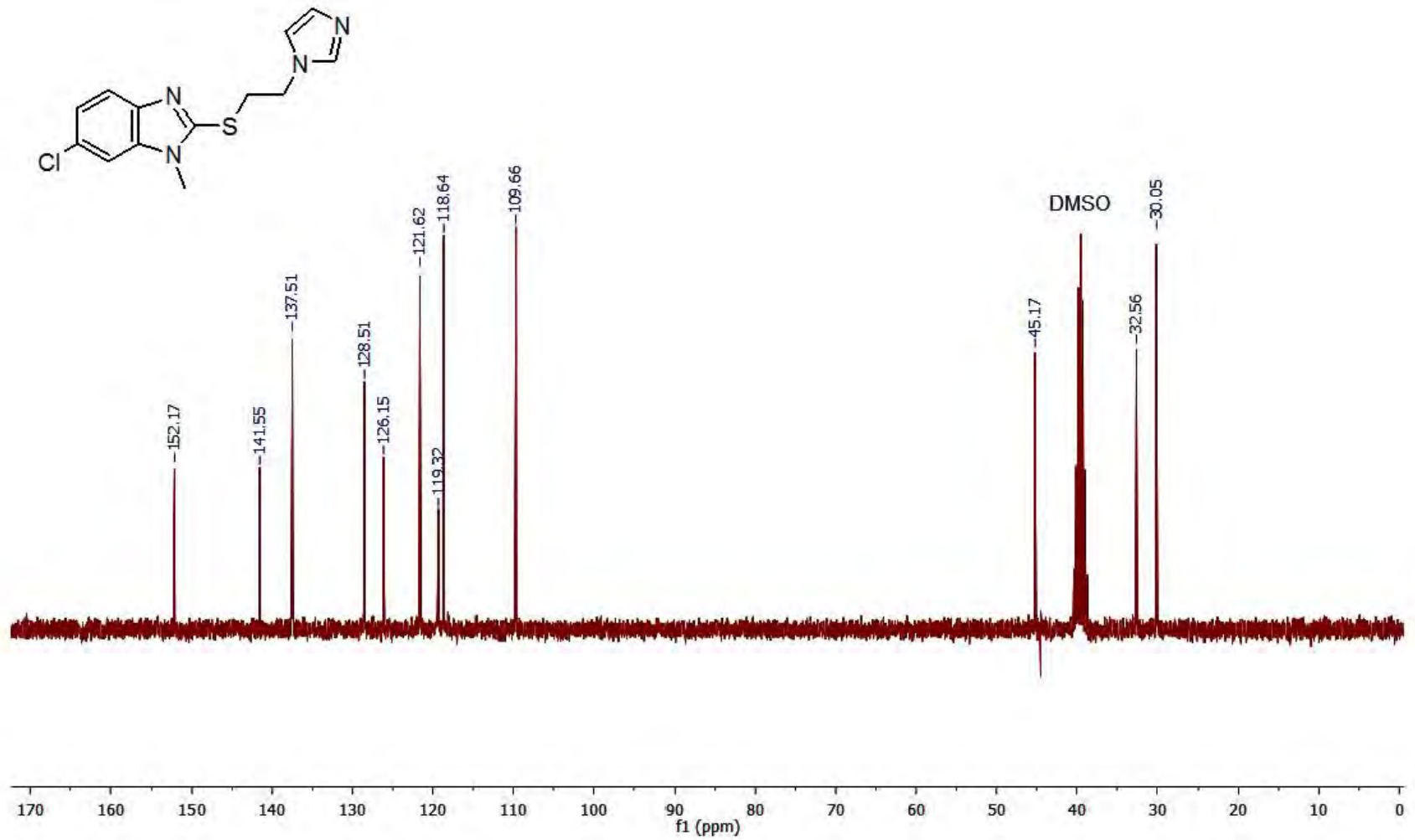
Espectro 10. RMN ^1H (300 MHz) del 2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1-metil-1*H*-bencimidazol (**S1J04**)

Espectro 11. RMN ^{13}C (75.5 MHz) del 2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1-metil-1H-bencimidazol (S1J04)



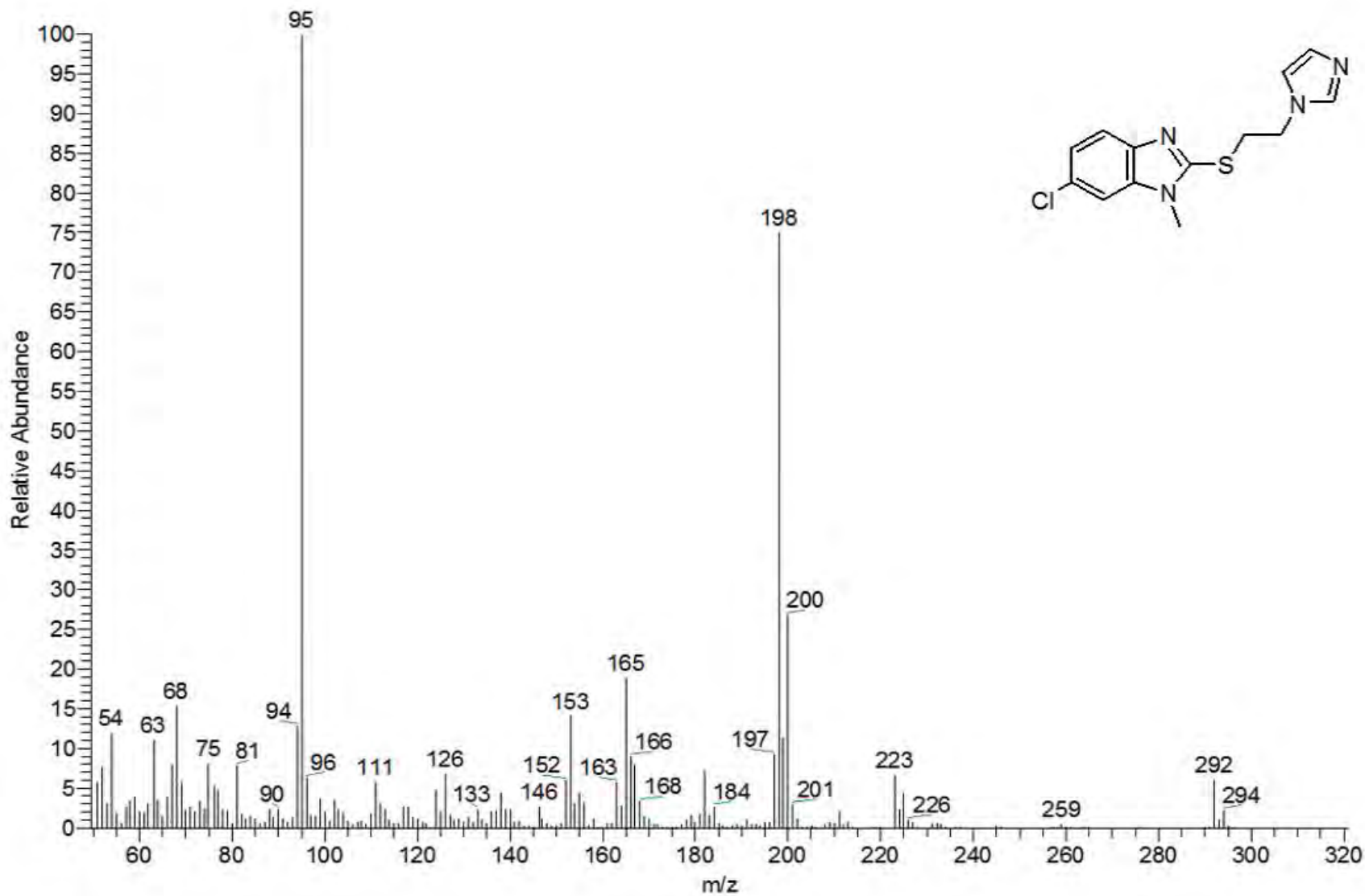
Espectro 12. EM (IE) del 2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1-metil-1H-bencimidazol (S1J04)

Espectro 13. RMN ^1H (300 MHz) del 6-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S1J05)

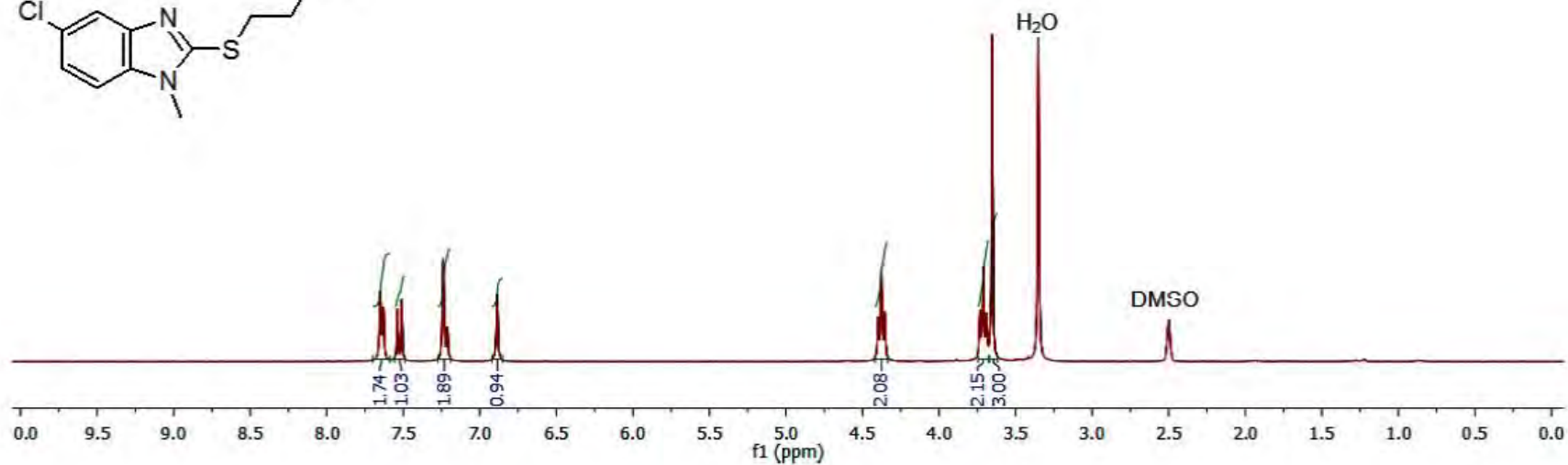
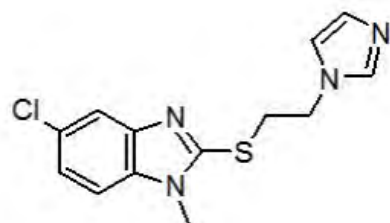
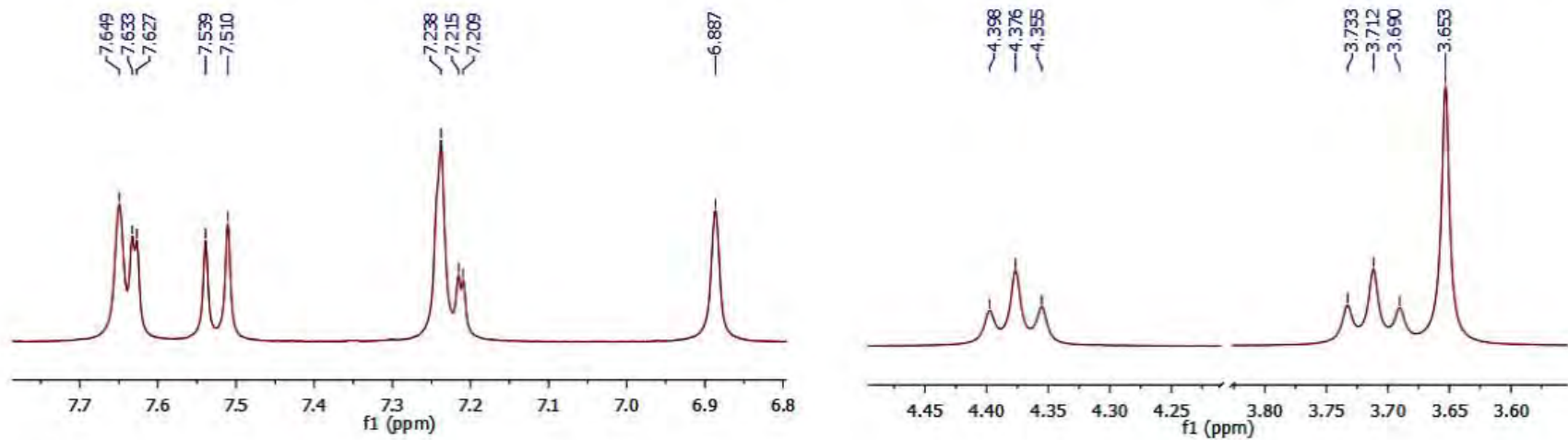


Espectro 14. RMN ¹³C (75.5 MHz) del 6-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S1J05)

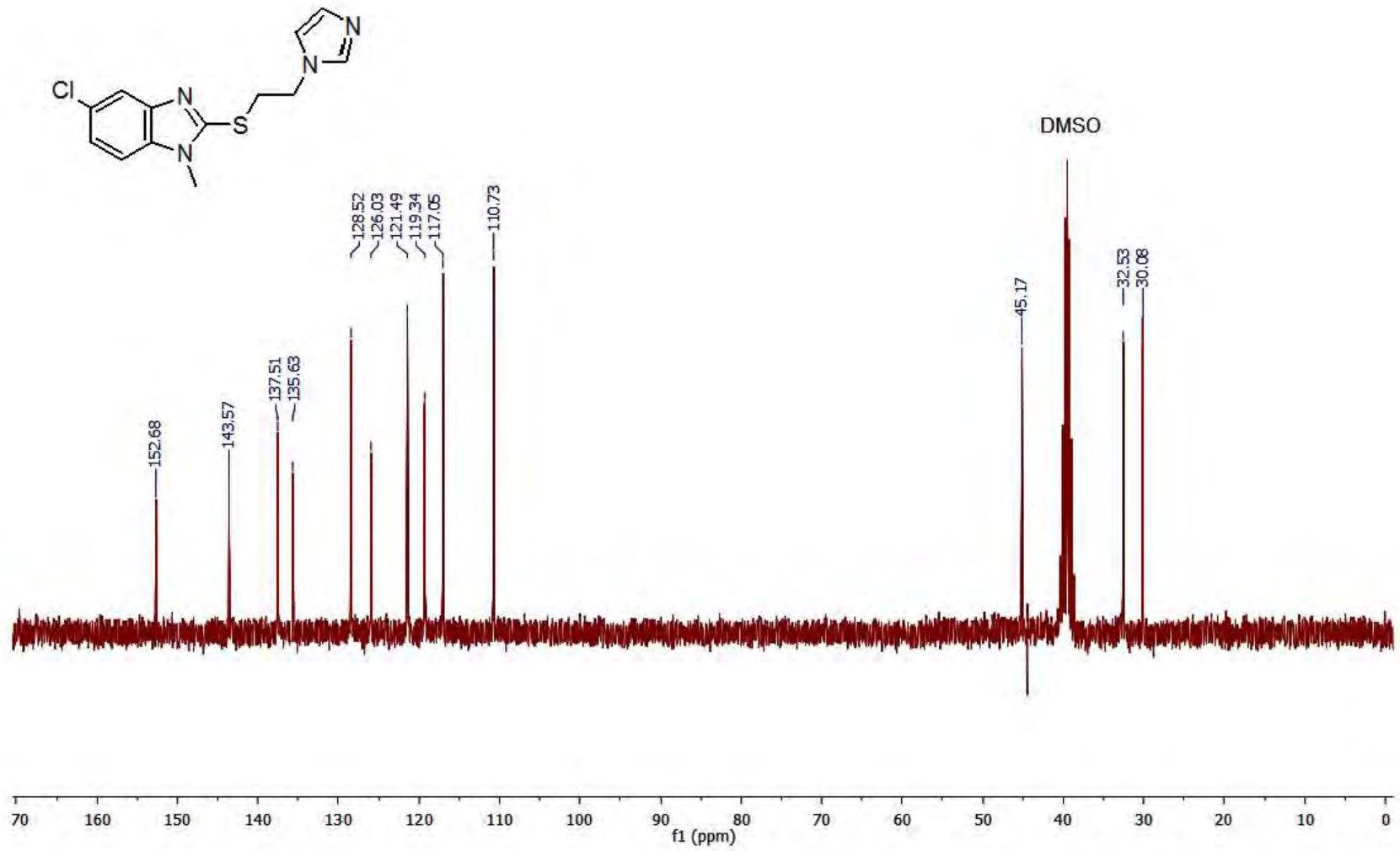
180



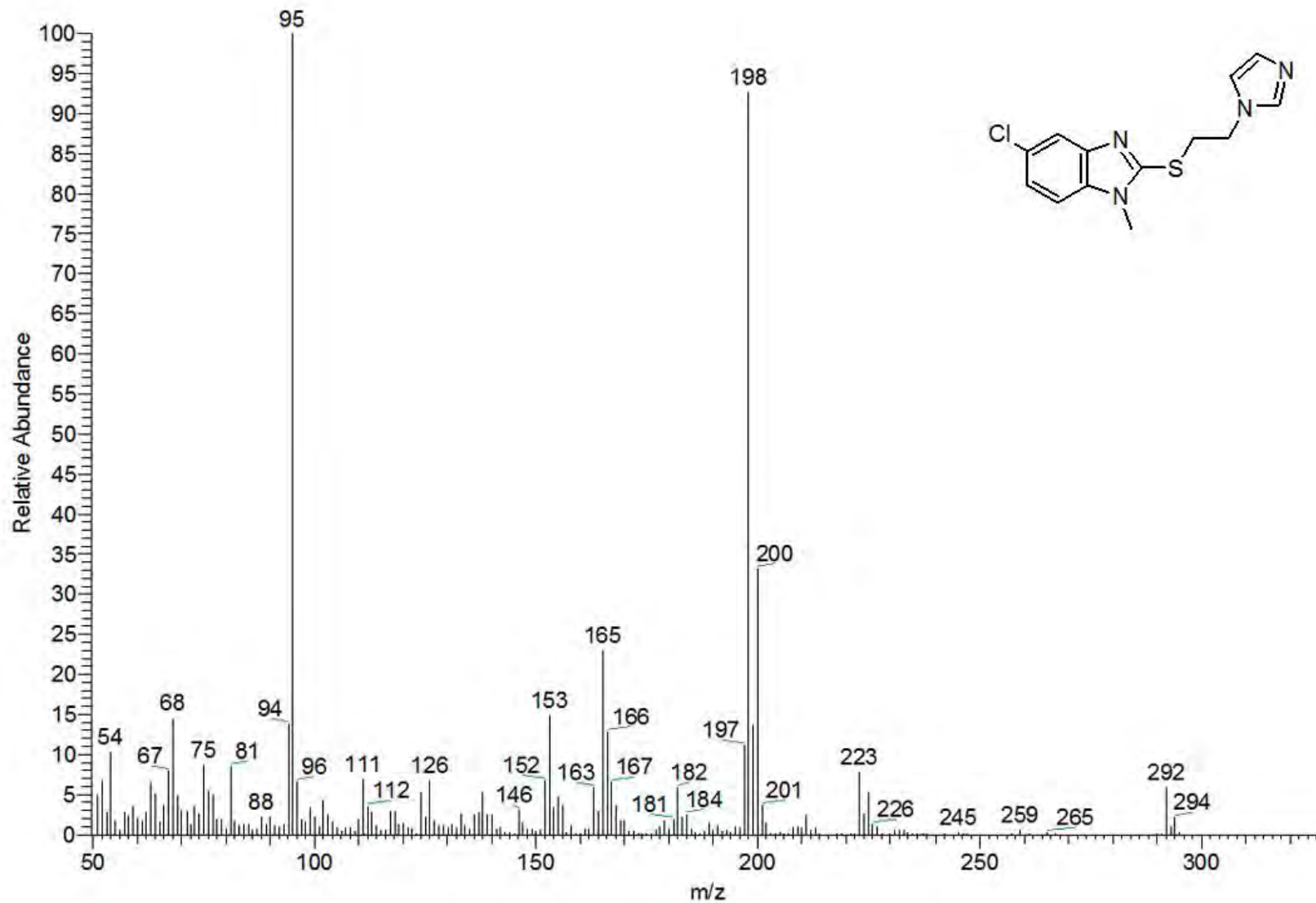
Espectro 15. EM (IE) del 6-cloro-2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1-metil-1H-bencimidazol (S1J05)



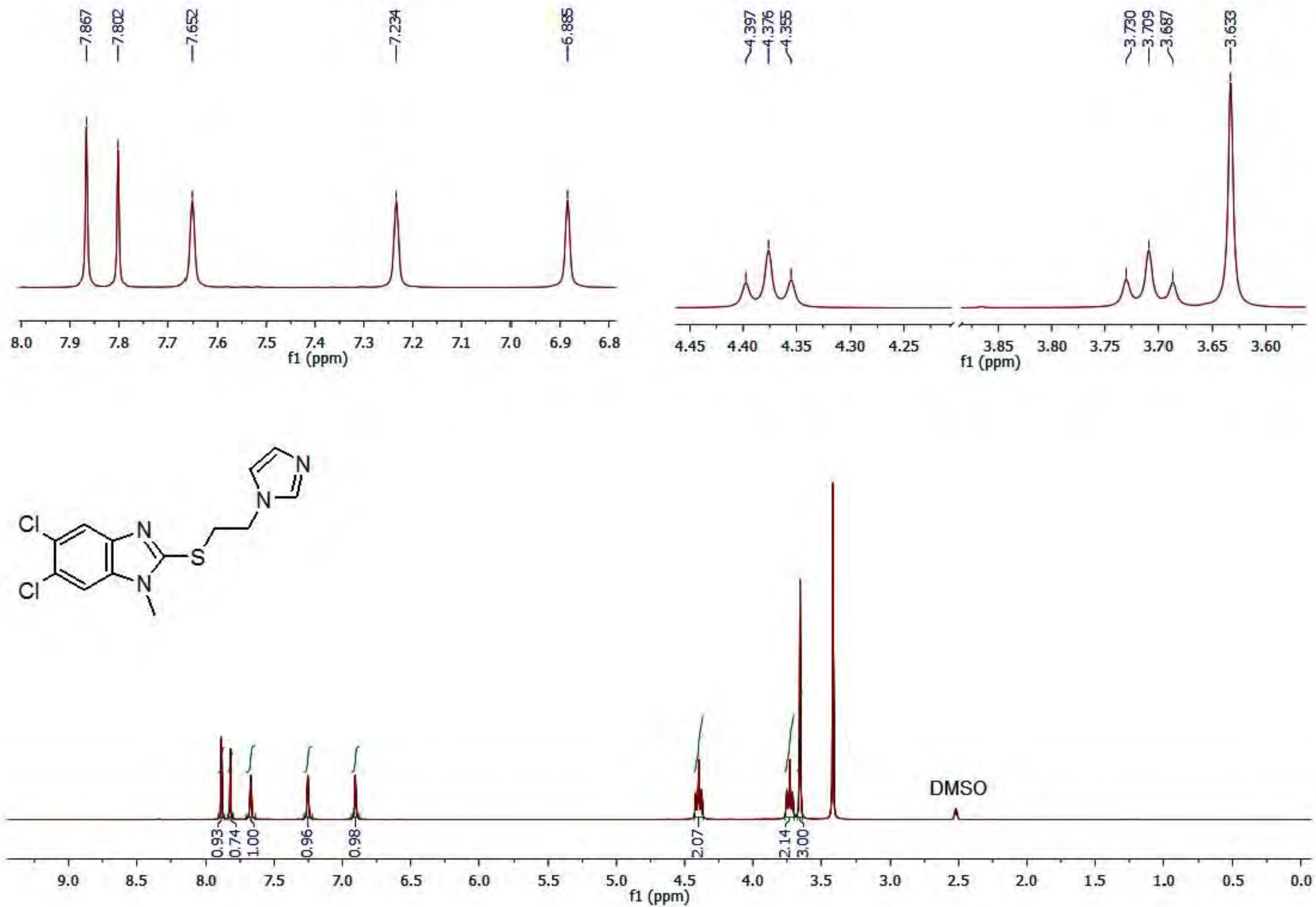
Espectro 16. RMN ^1H (300 MHz) del 5-cloro-2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1-metil-1*H*-bencimidazol (**S1J06**)



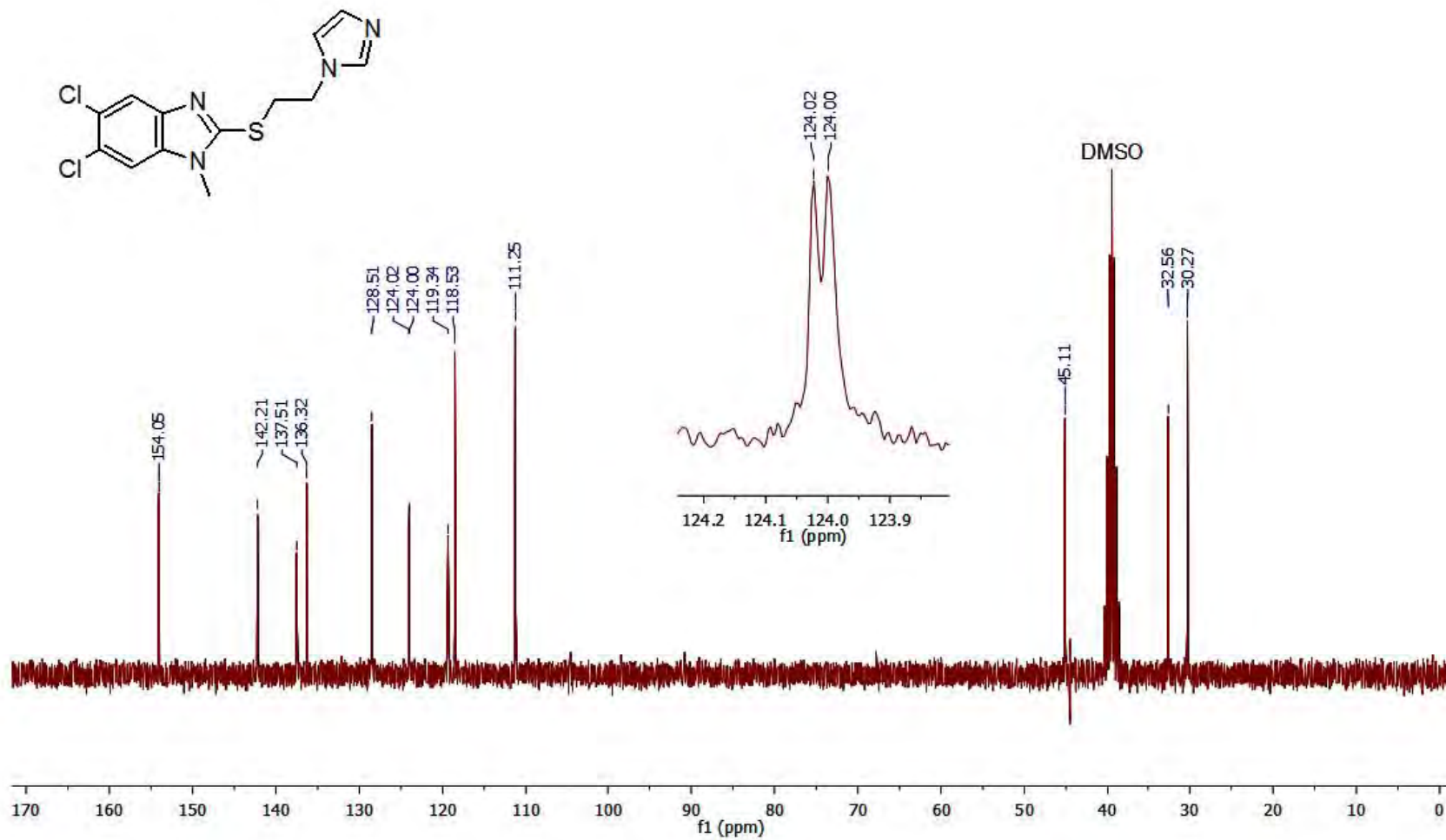
Espectro 17. RMN ¹³C (75.5 MHz) del 5-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S1J06)

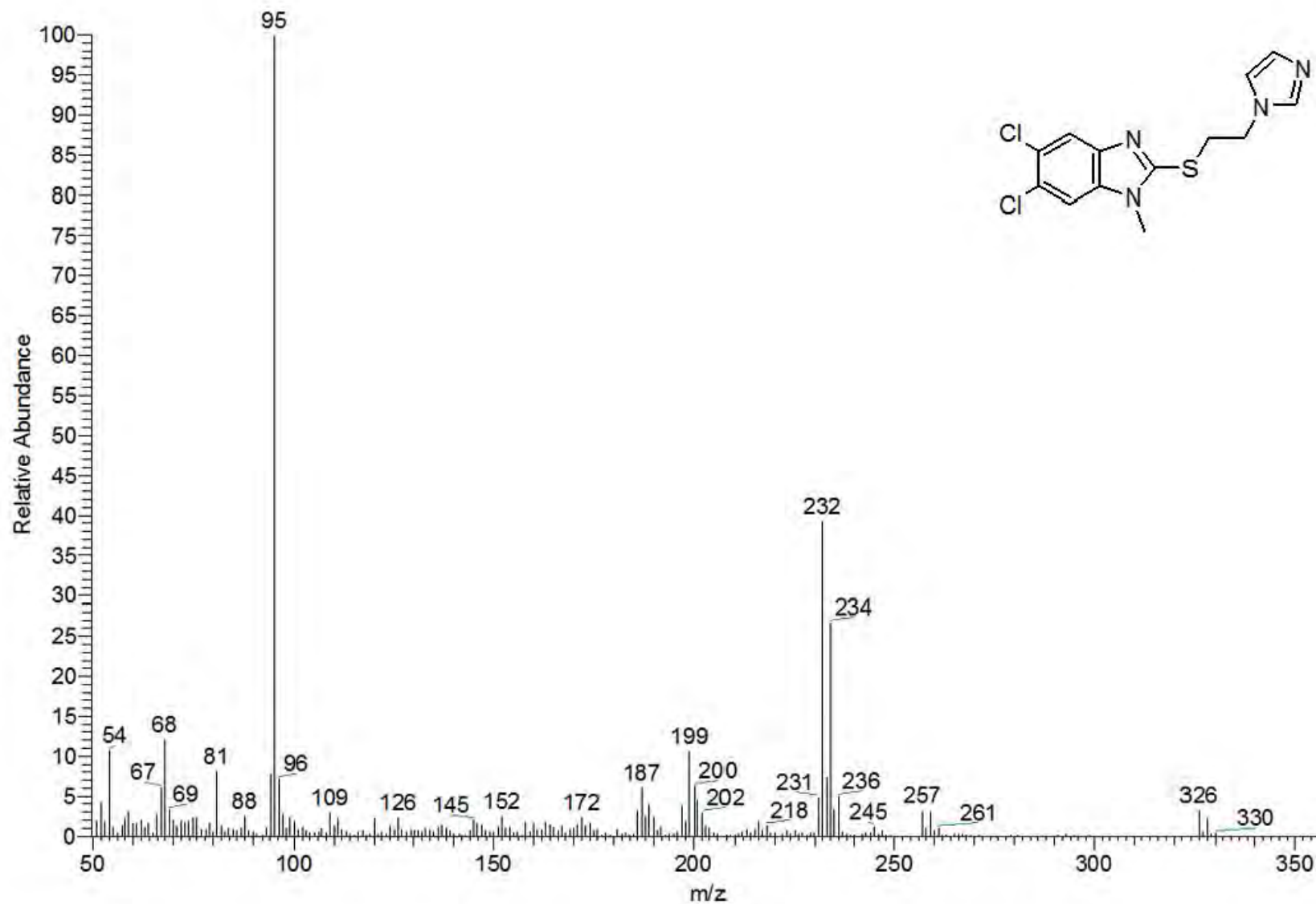


Espectro 18. EM (IE) del 5-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S1J06)

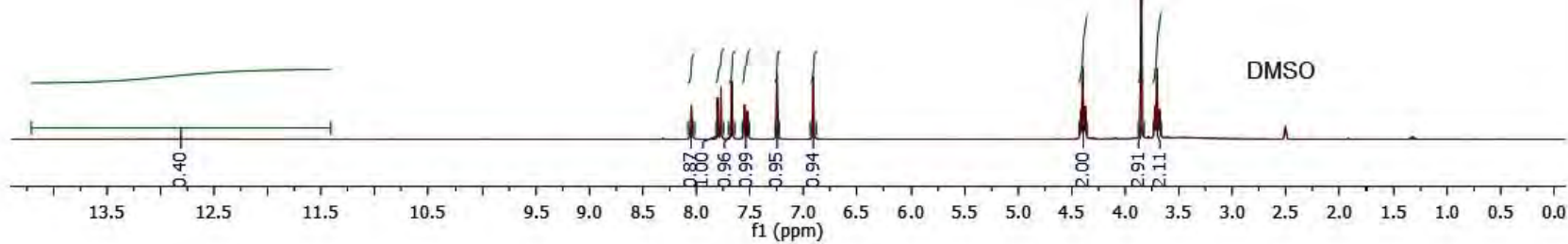
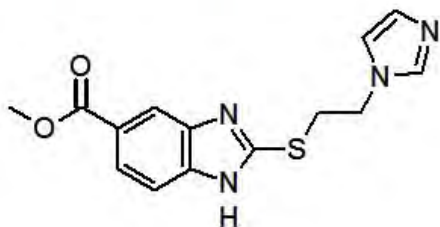
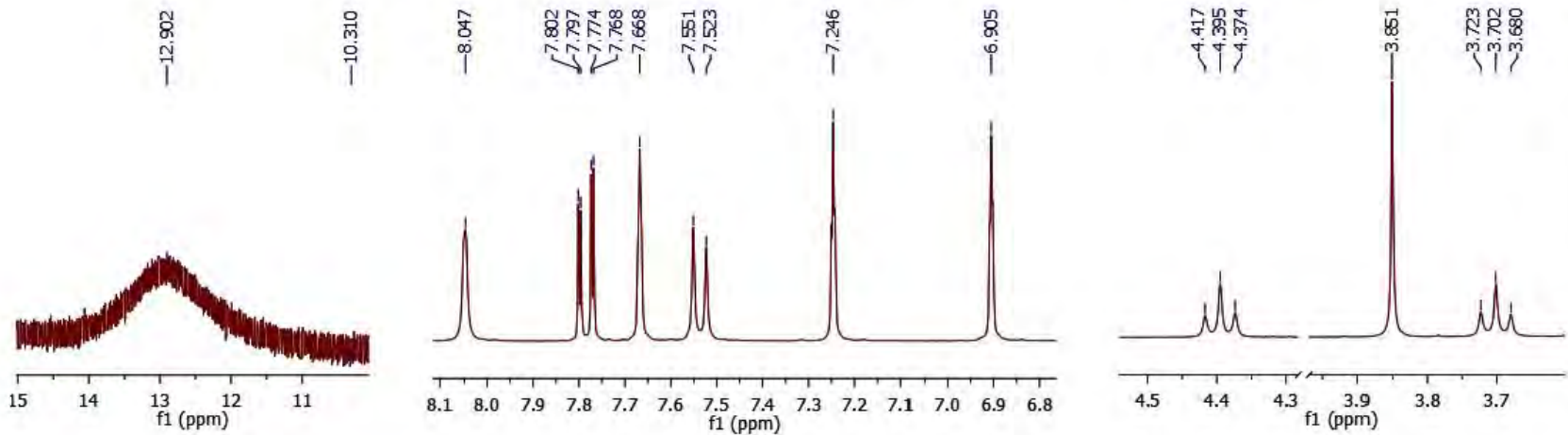


Espectro 19. RMN ^1H (300 MHz) del 5,6-dicloro-2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1-metil-1H-bencimidazol (S1J07)

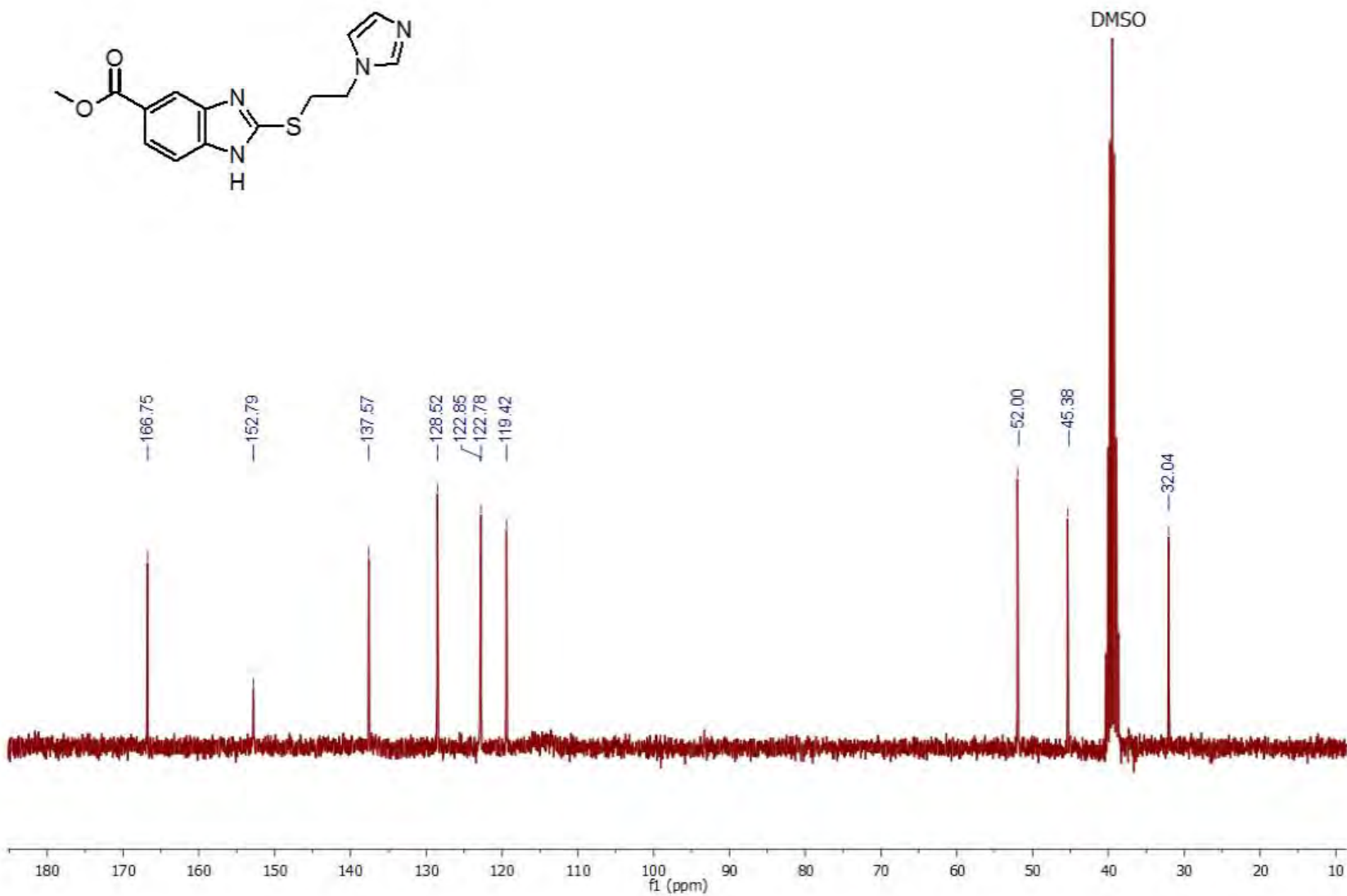
Espectro 20. RMN ^{13}C (75.5 MHz) del 5,6-dicloro-2-[(1H-imidazol-1-il)etil]sulfanil)-1-metil-1H-bencimidazol (S1J07)

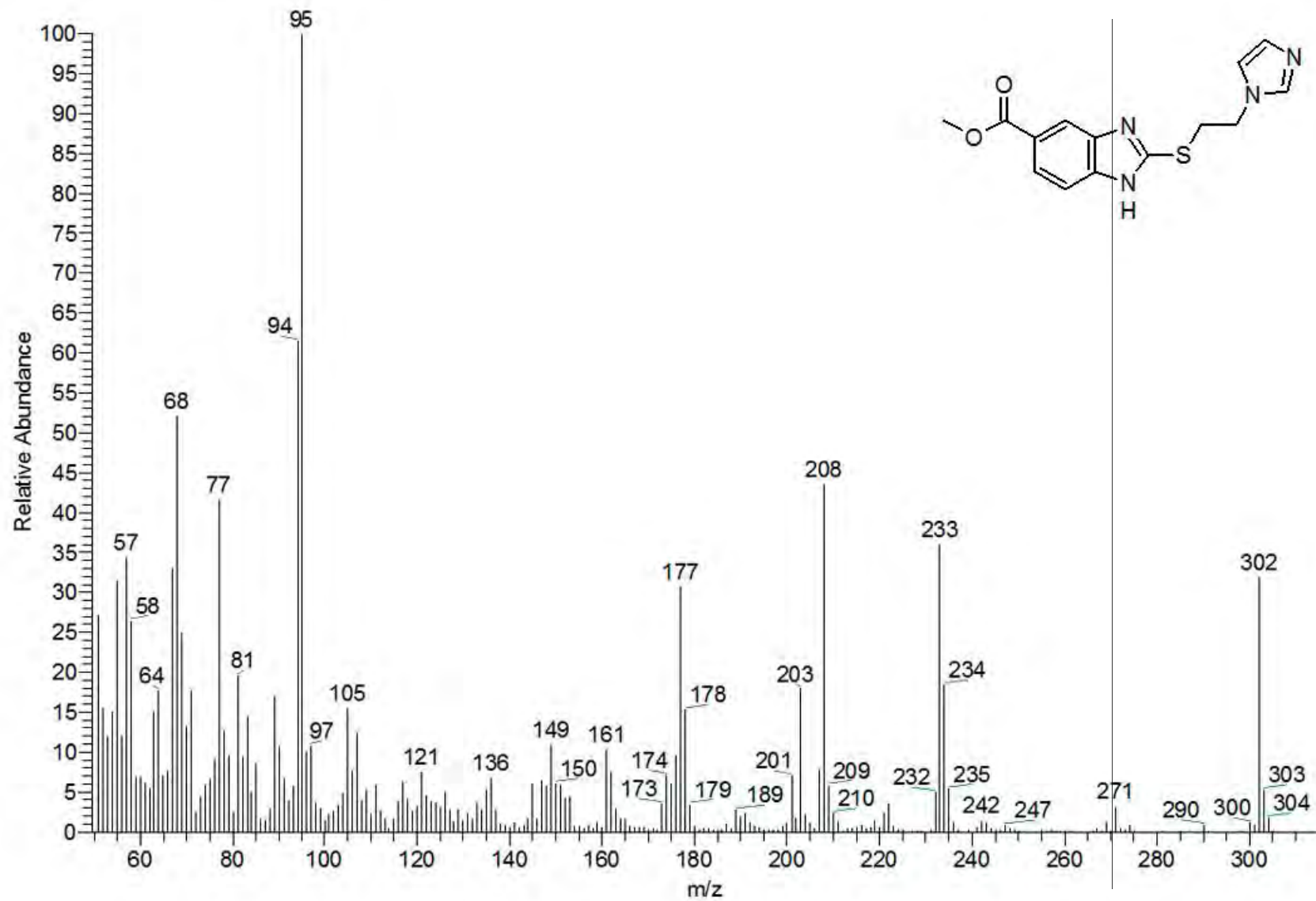


Espectro 21. EM (IE) del 5,6-dicloro-2-([2-(1H-imidazol-1-yl)etil]sulfanil)-1-metil-1H-bencimidazol (S1J07)

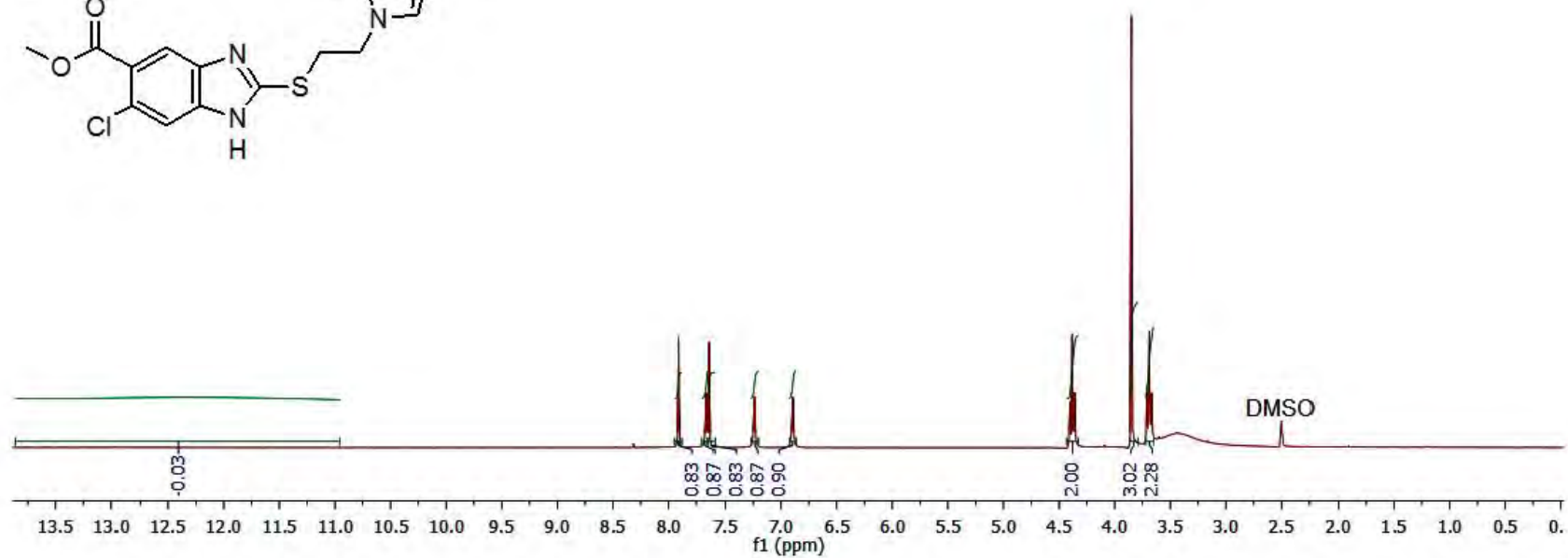
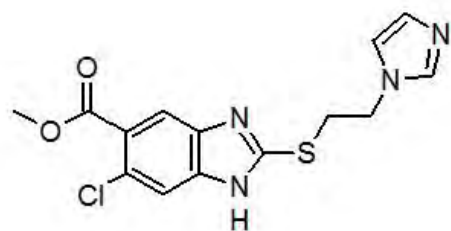
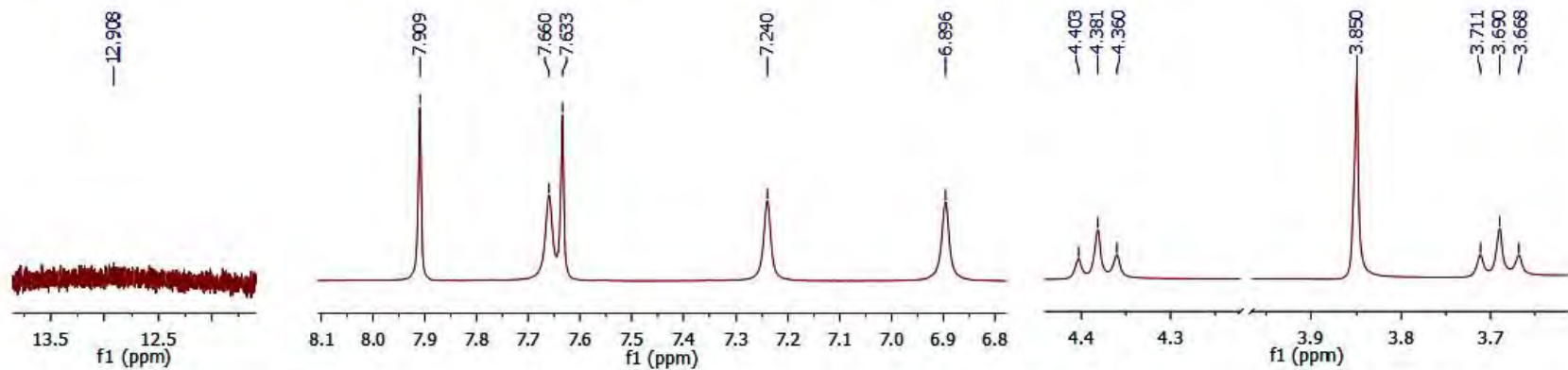


Espectro 22. RMN ^1H (300 MHz) del 2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1*H*-benzimidazol-5(6)-carboxilato de metilo (**S2J08**)

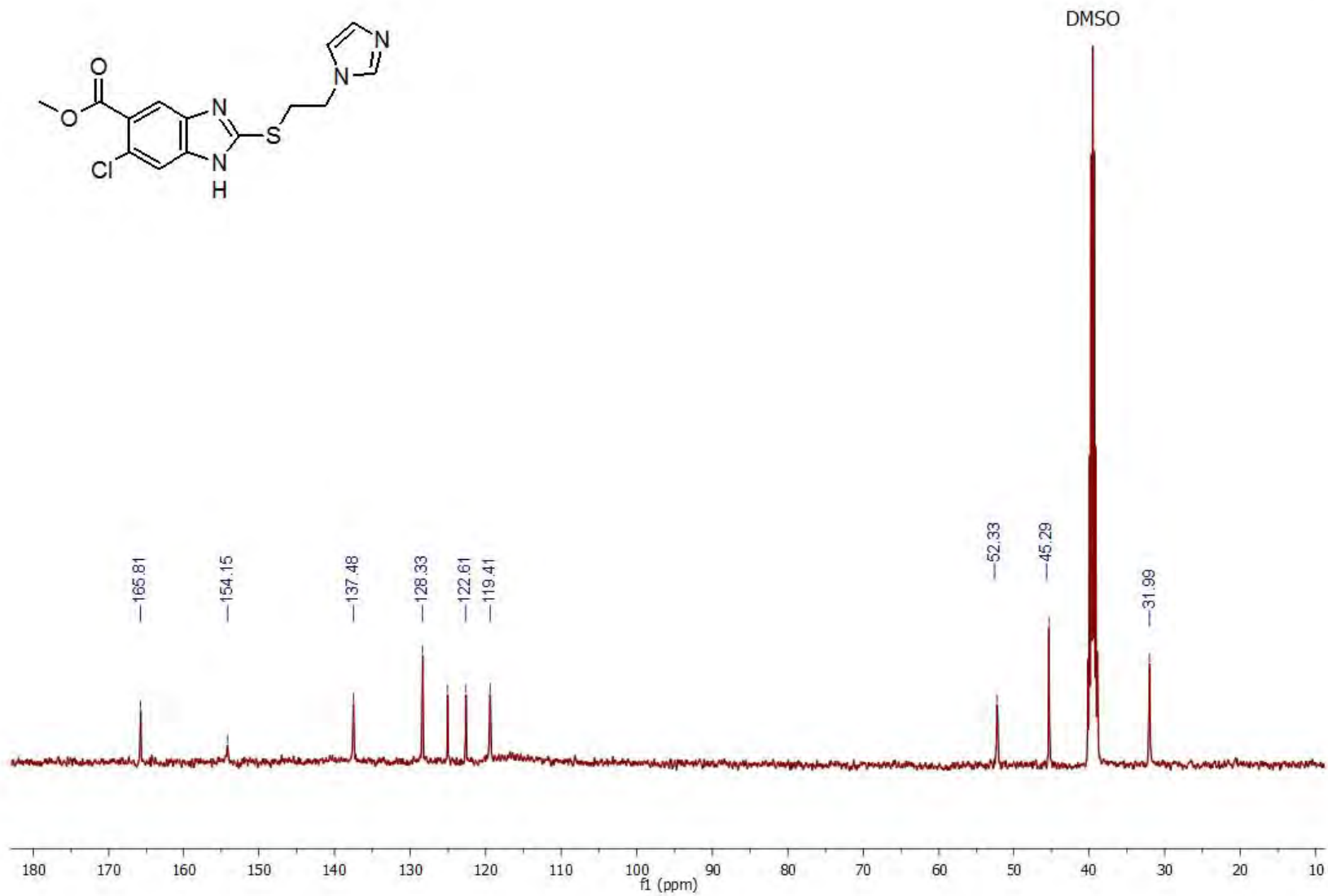
Espectro 23. RMN ¹³C (75.5 MHz) del 2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1H-bencimidazol-5(6)-carboxilato de metilo (S2J08)

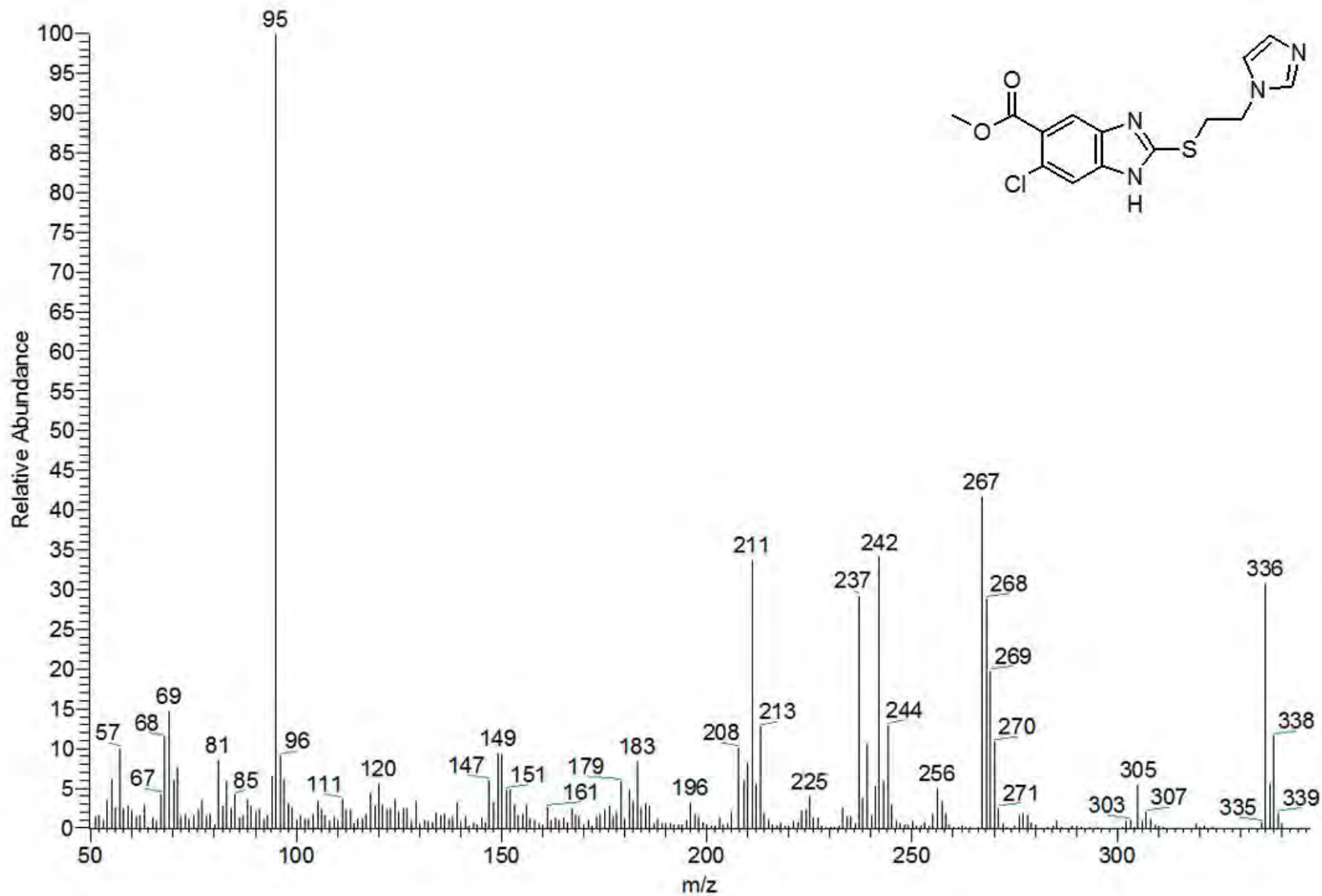


Espectro 24. EM (IE) del 2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1H-benzimidazol-5(6)-carboxilato de metilo (S2J08)

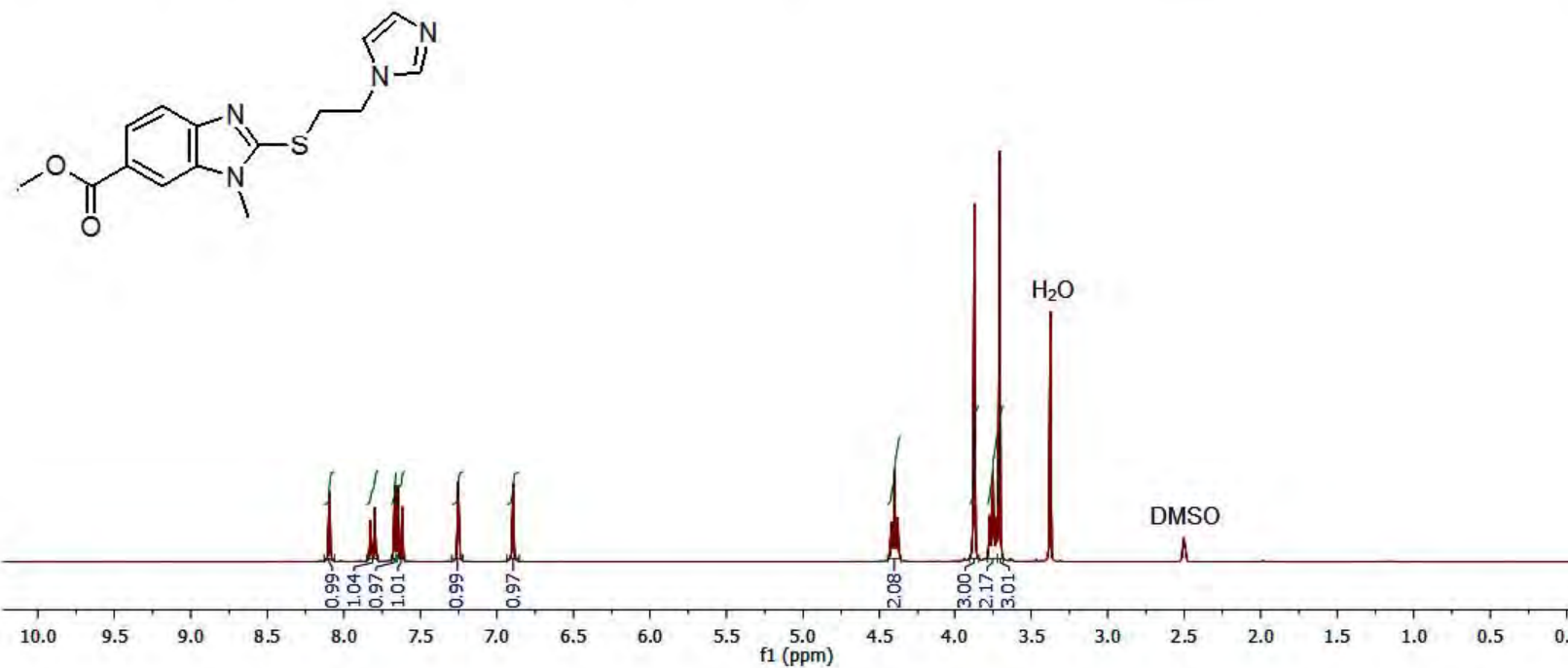
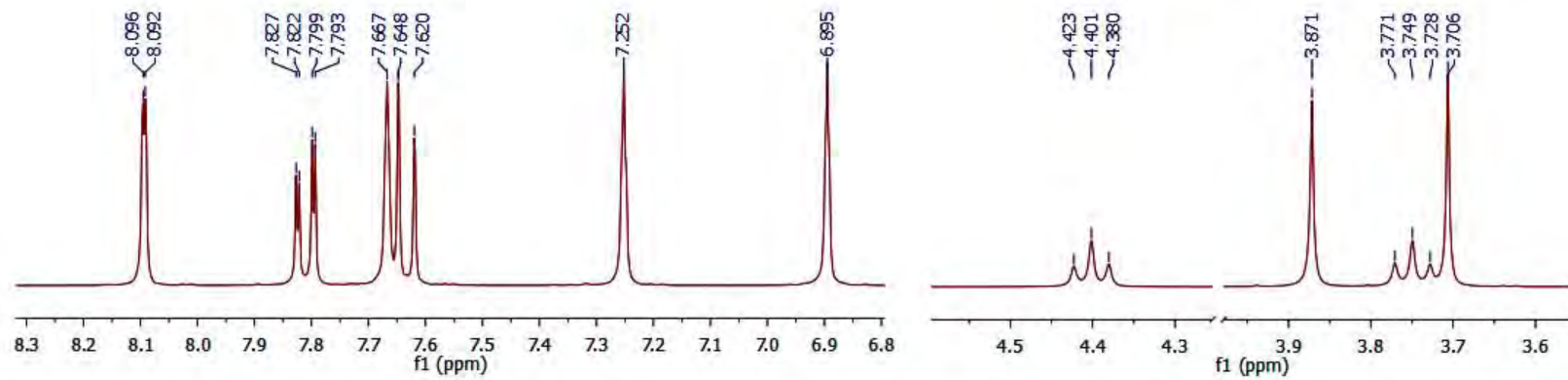


Espectro 25. RMN ¹H (300 MHz) del 6-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-benzimidazol-5-carboxilato de metilo (S2J09)

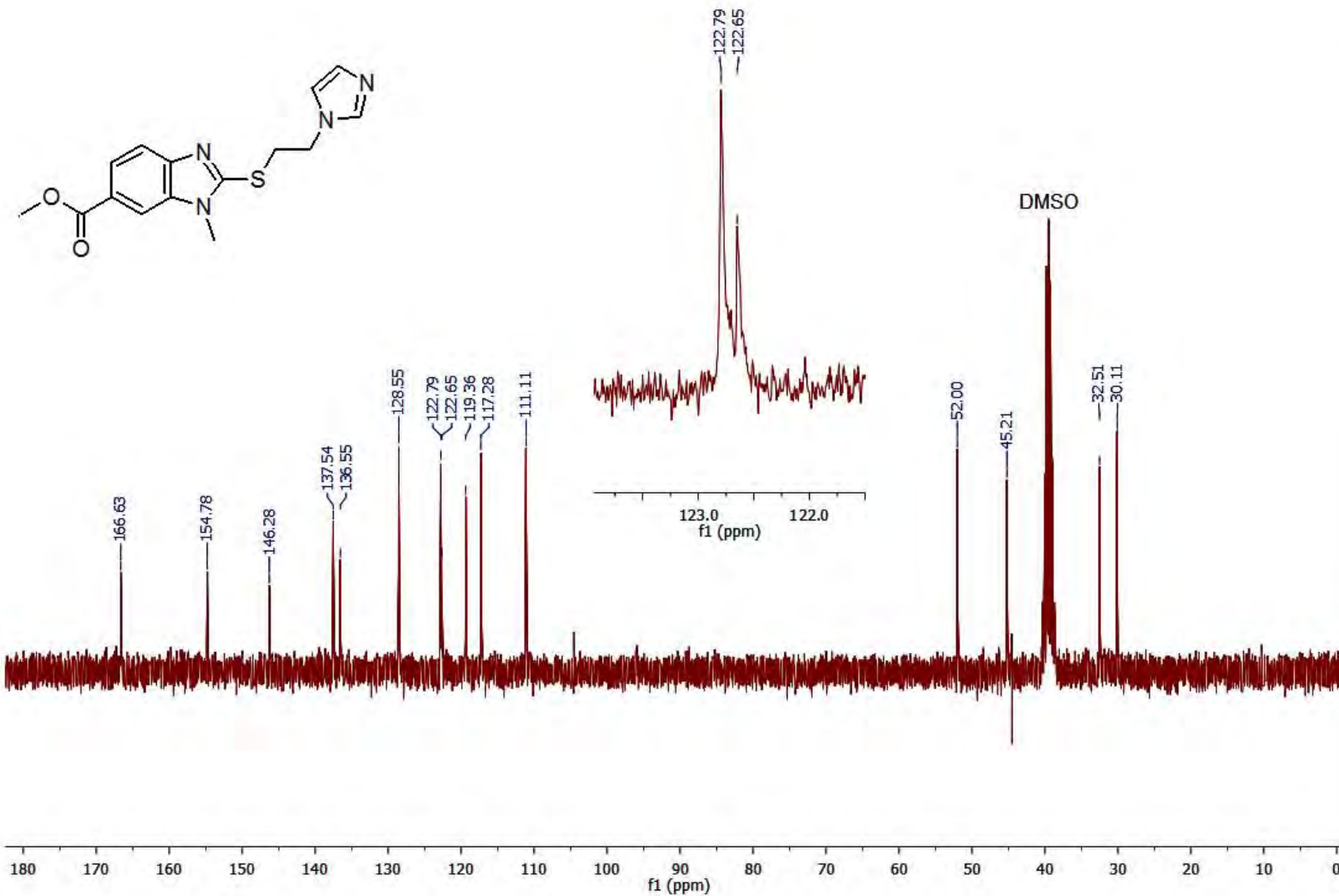
Espectro 26. RMN ^{13}C (75.5 MHz) del 6-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol-5-carboxilato de metilo (S2J09)

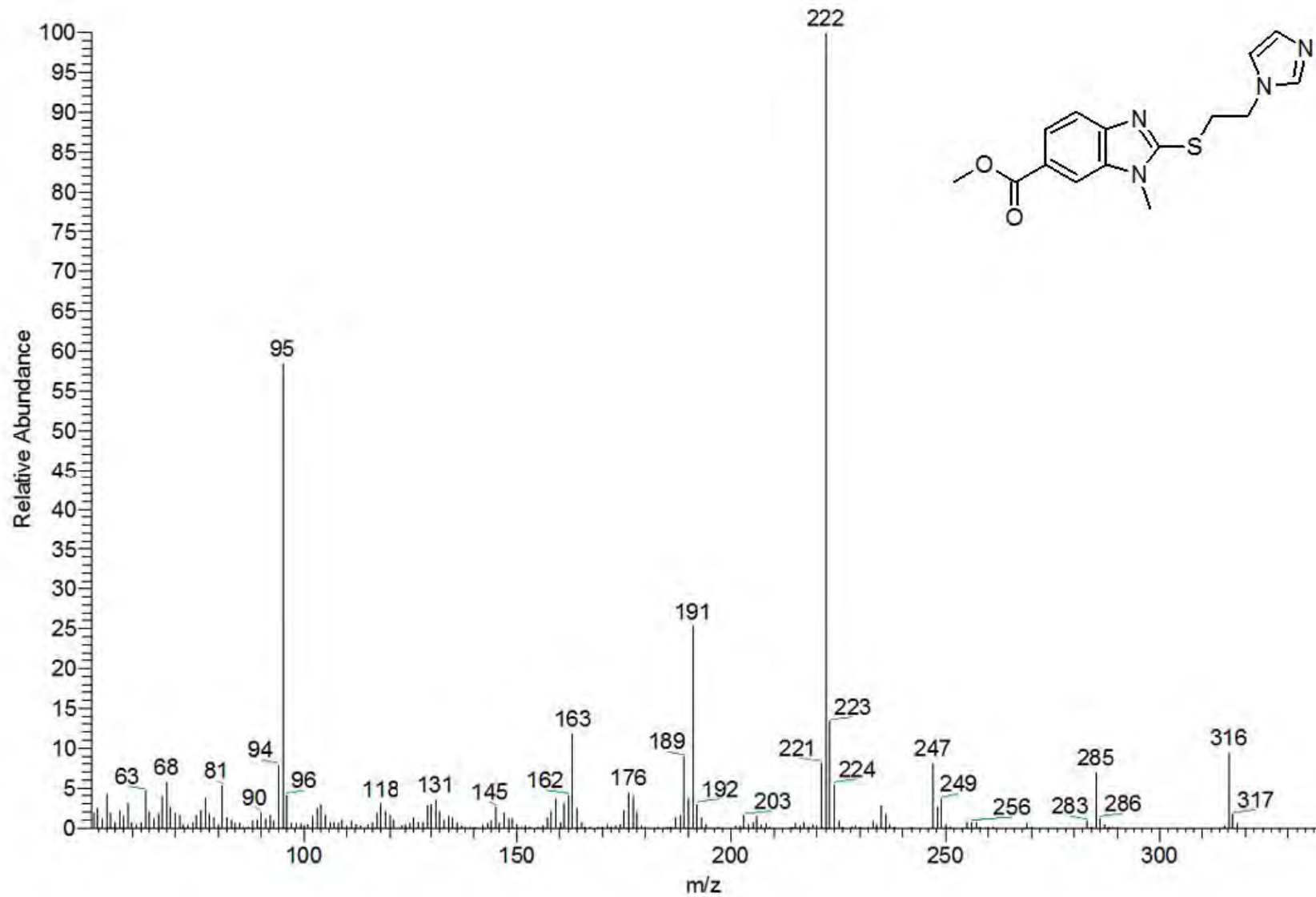


Espectro 27. EM (IE) del 6-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol-5-carboxilato de metilo (S2J09)

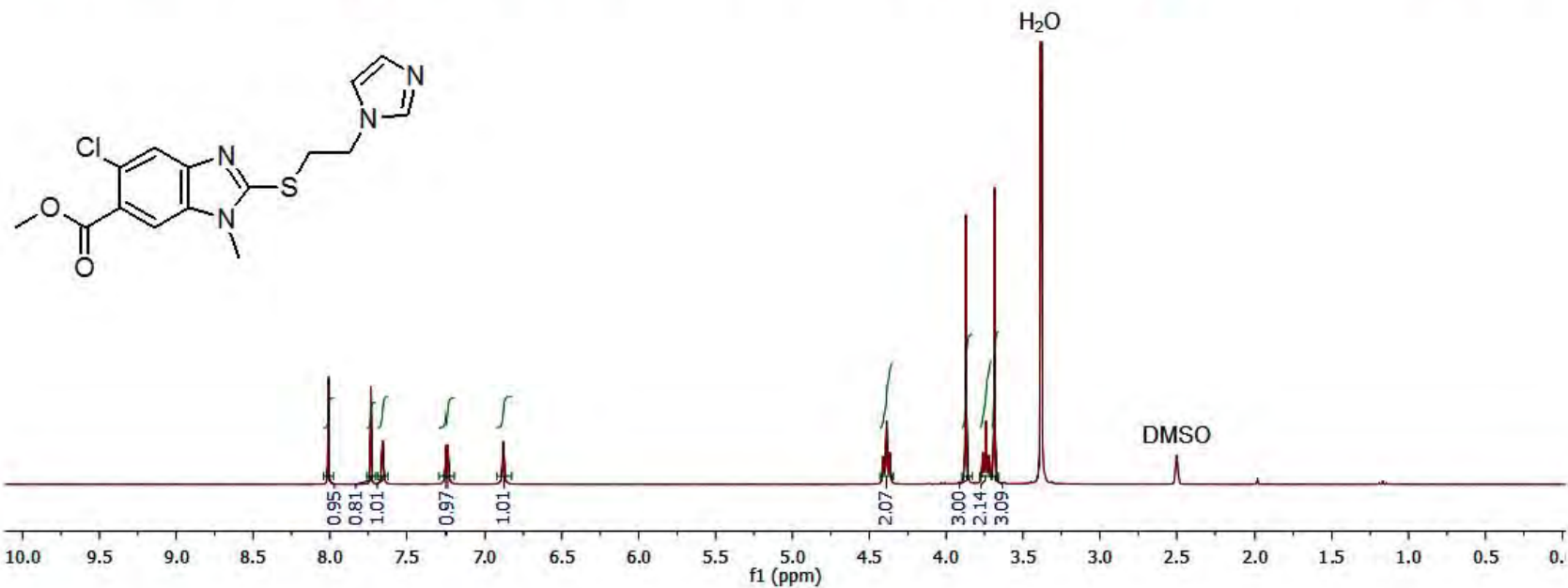
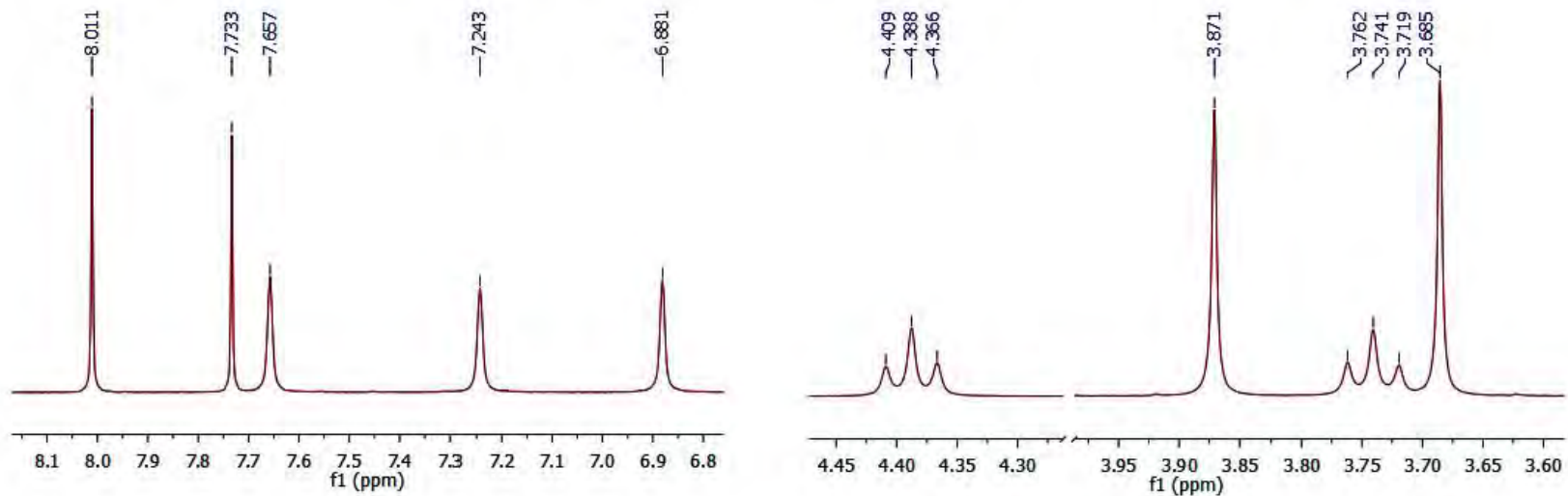


Espectro 28. RMN ^1H (300 MHz) del 2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-6-carboxilato de metilo (**S2J10**)

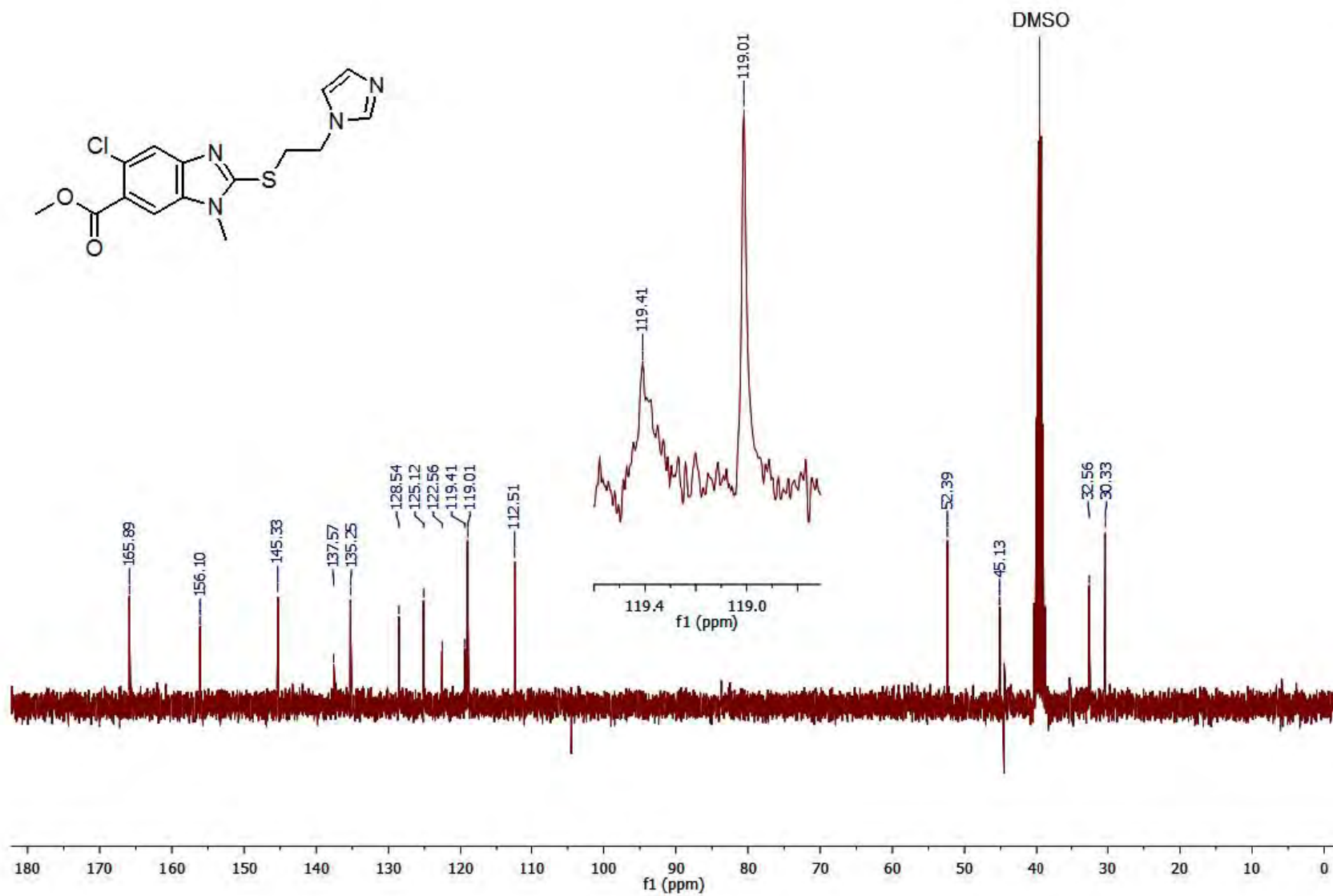
Espectro 29. RMN ^{13}C (75.5 MHz) del 2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-6-carboxilato de metilo (S2J10)



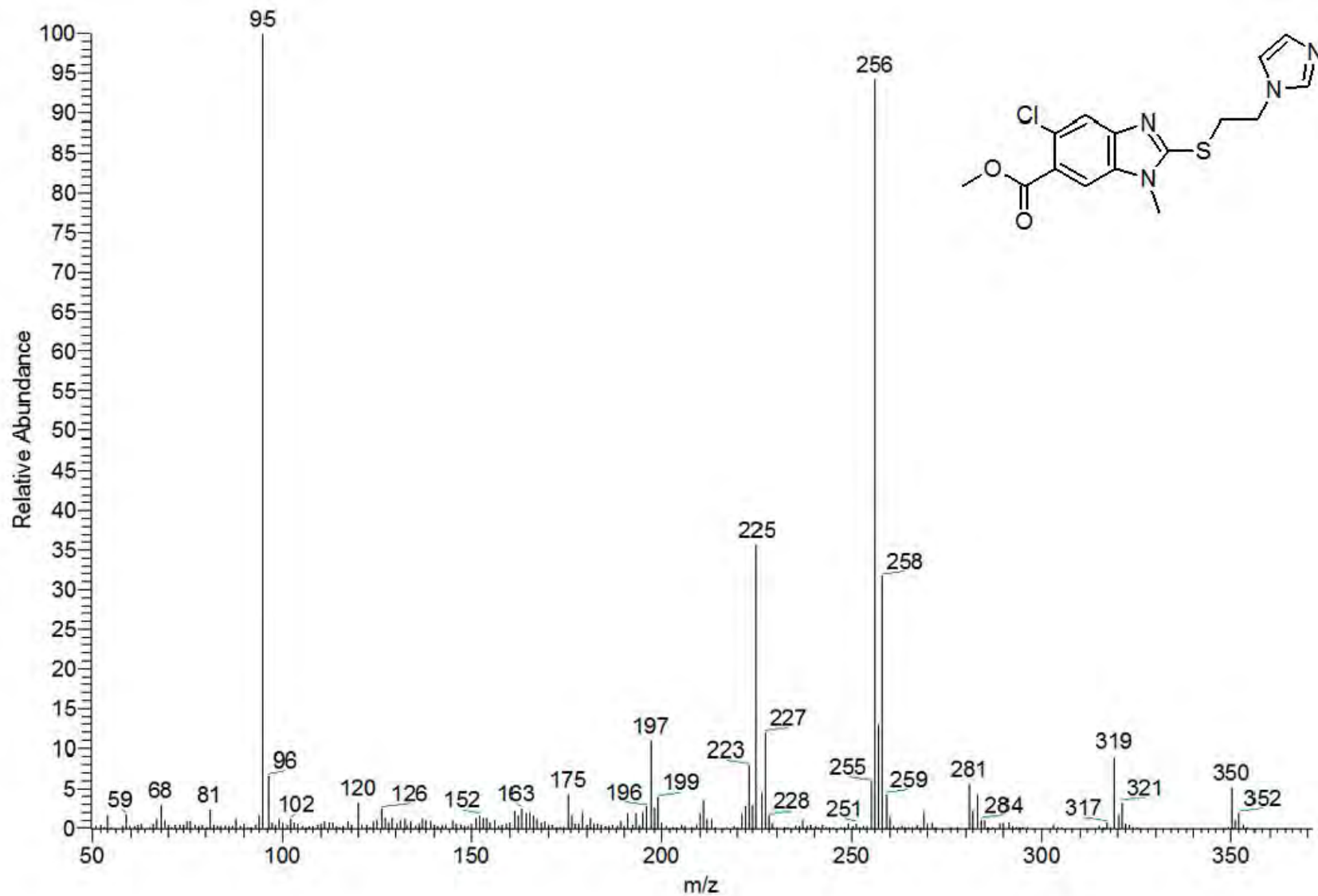
Espectro 30. EM (IE) del 2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1-metil-1H-bencimidazol-6-carboxilato de metilo (S2J10)

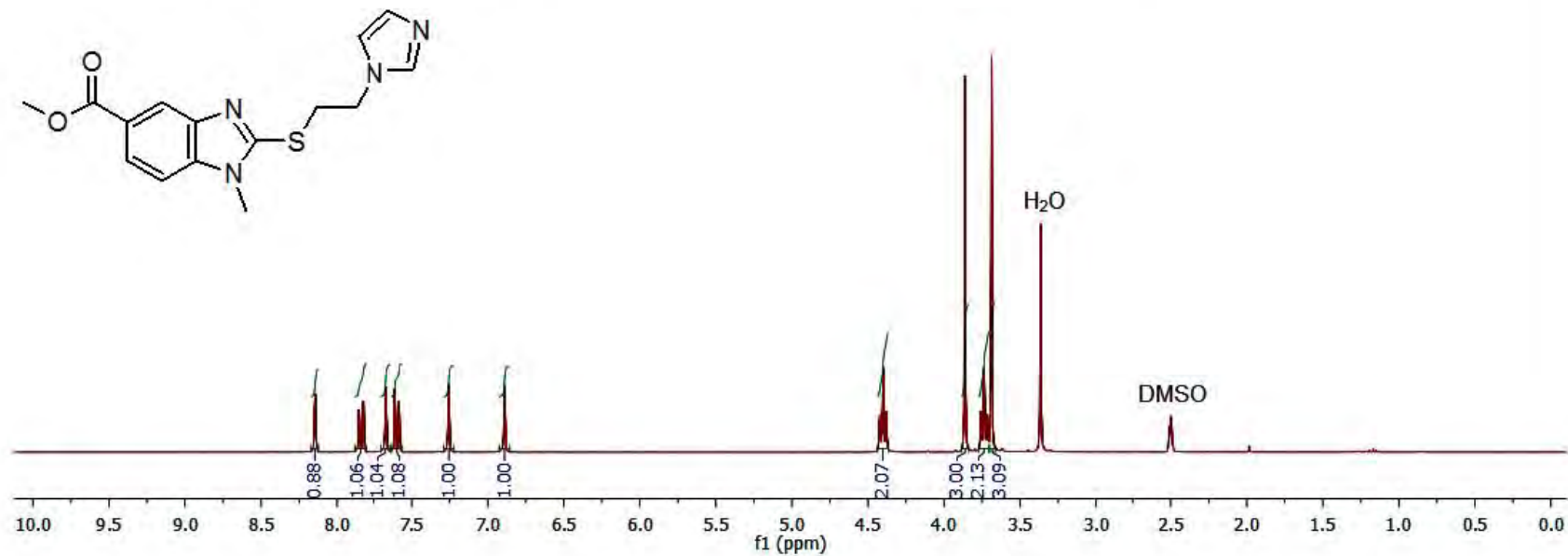
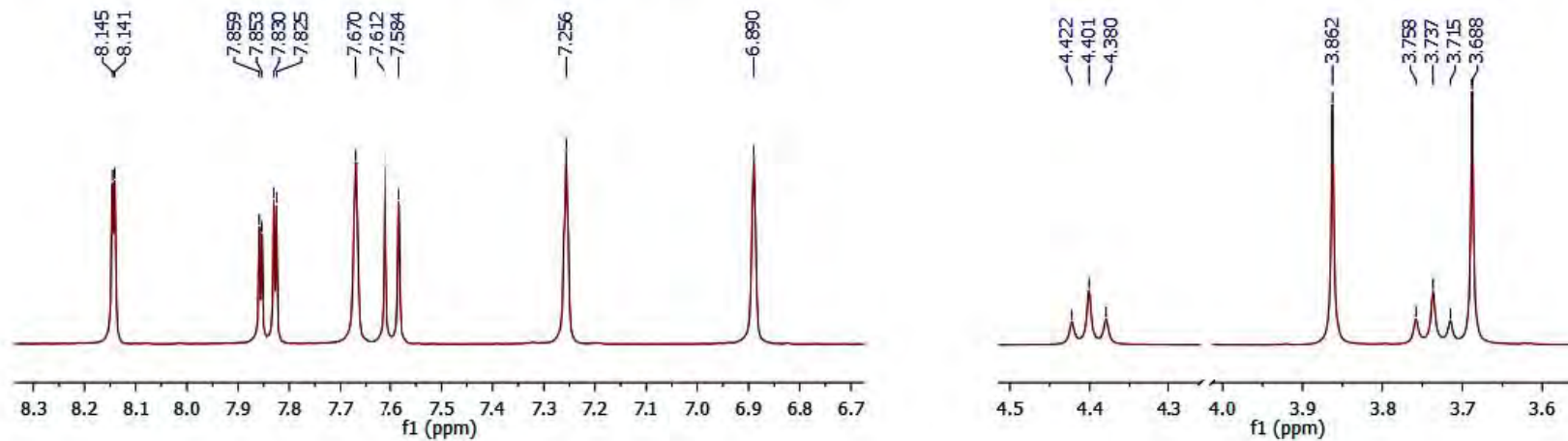


Espectro 31. RMN ^1H (300 MHz) del 5-cloro-2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1-metil-1*H*-bencimidazol-6-carboxilato de metilo (S2J11)

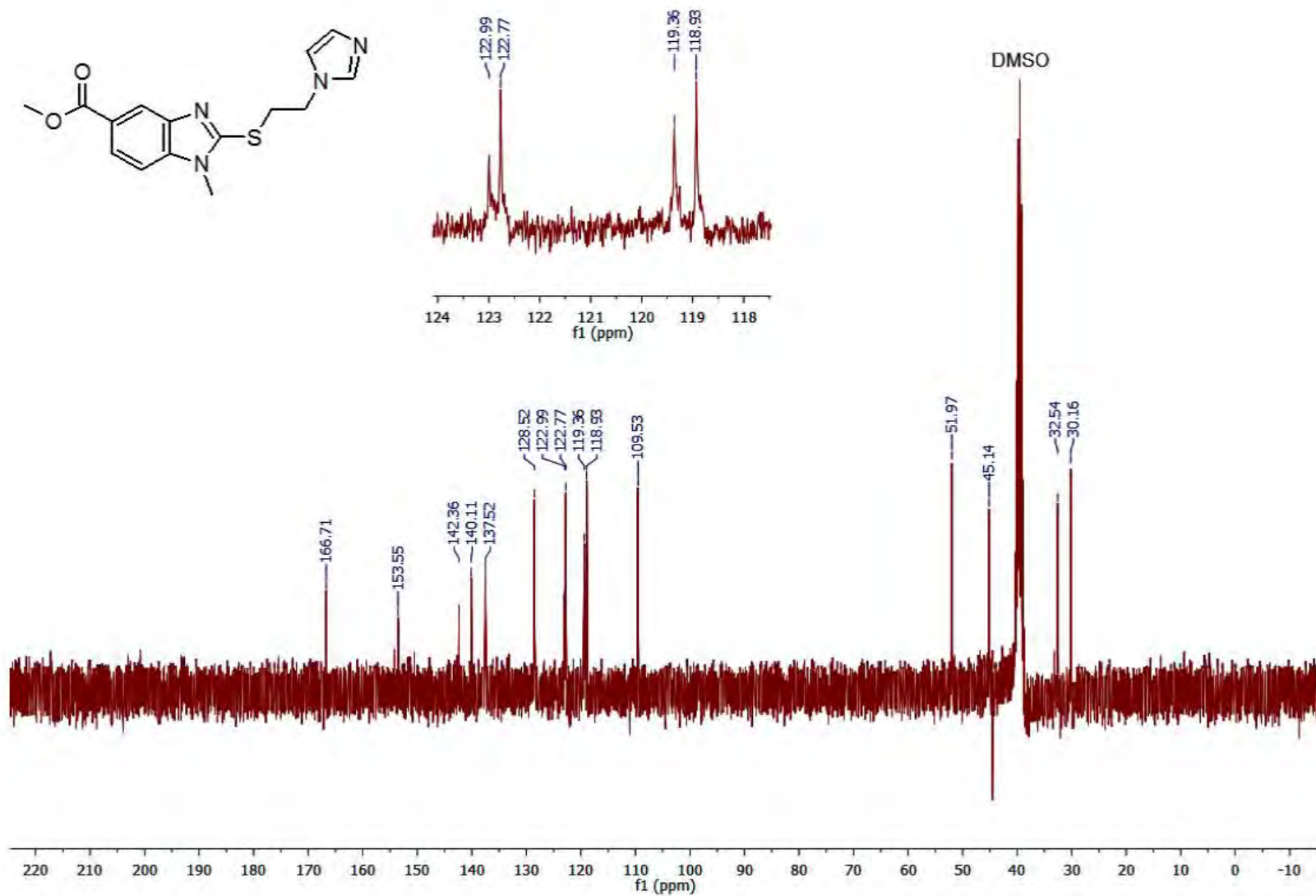


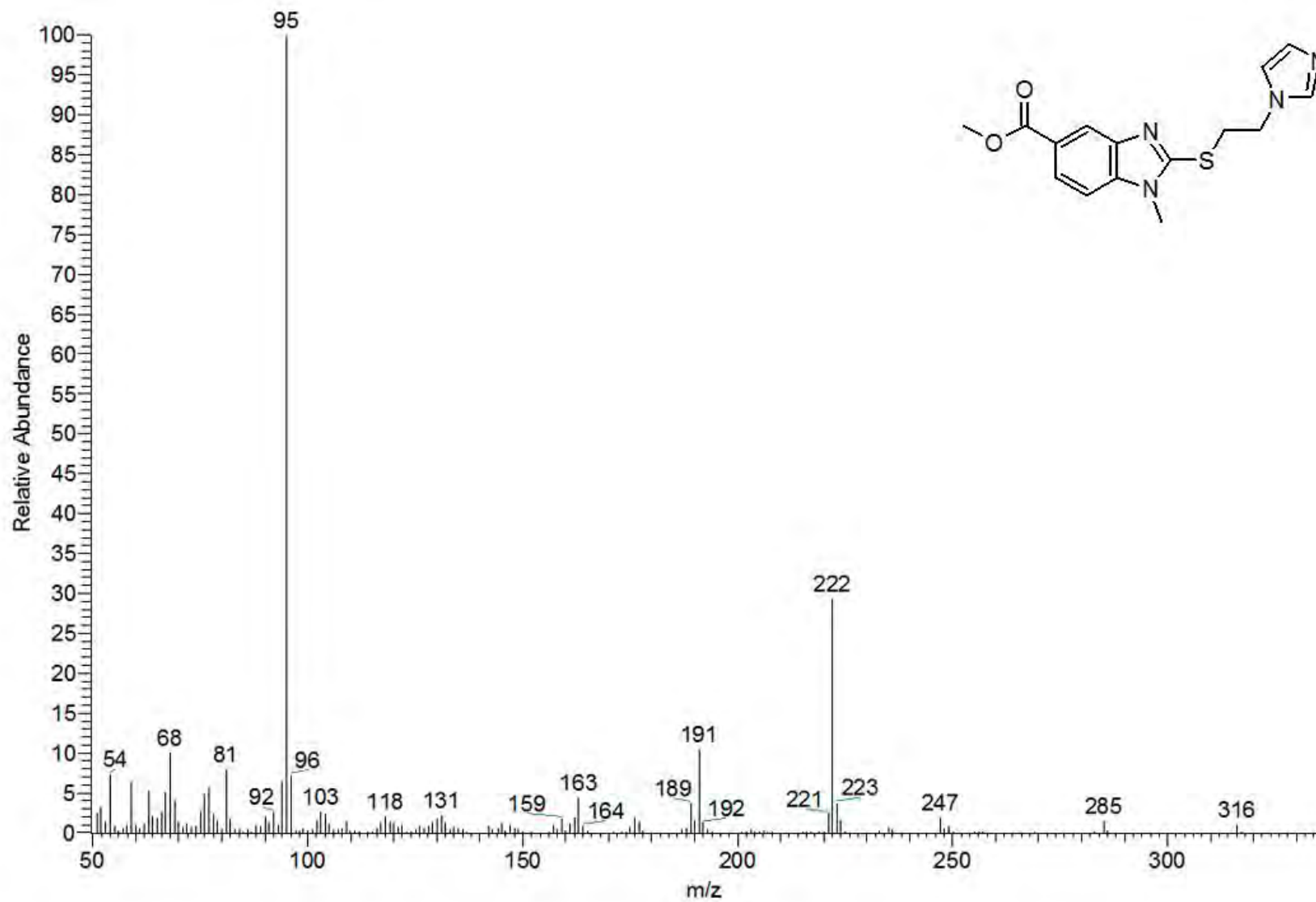
Espectro 32. RMN ^{13}C (75.5 MHz) del 5-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-6-carboxilato de metilo (S2J11)

Espectro 33. EM (IE) del 5-cloro-2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1-metil-1*H*-bencimidazol-6-carboxilato de metilo (**S2J11**)

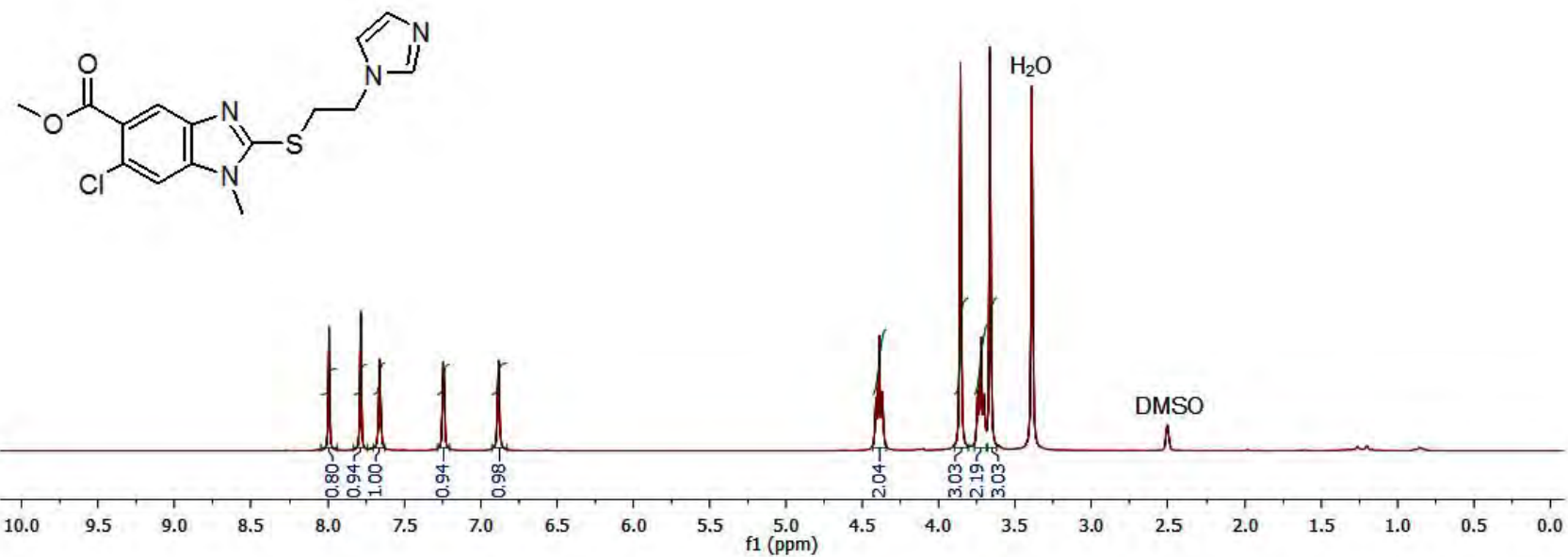
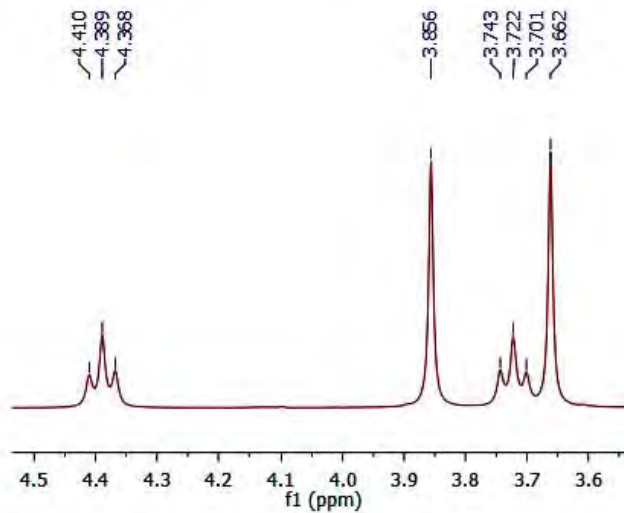
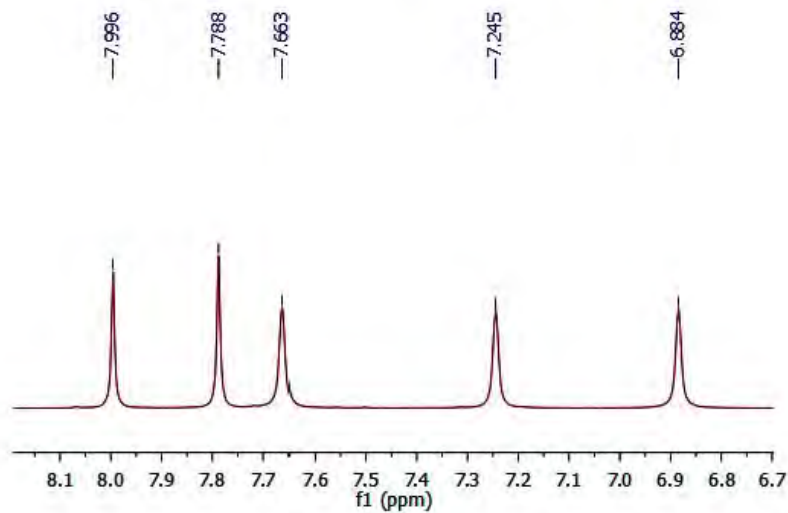


Espectro 34. RMN ^1H (300 MHz) del 2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1-metil-1*H*-bencimidazol-5-carboxilato de metilo (**S2J12**)

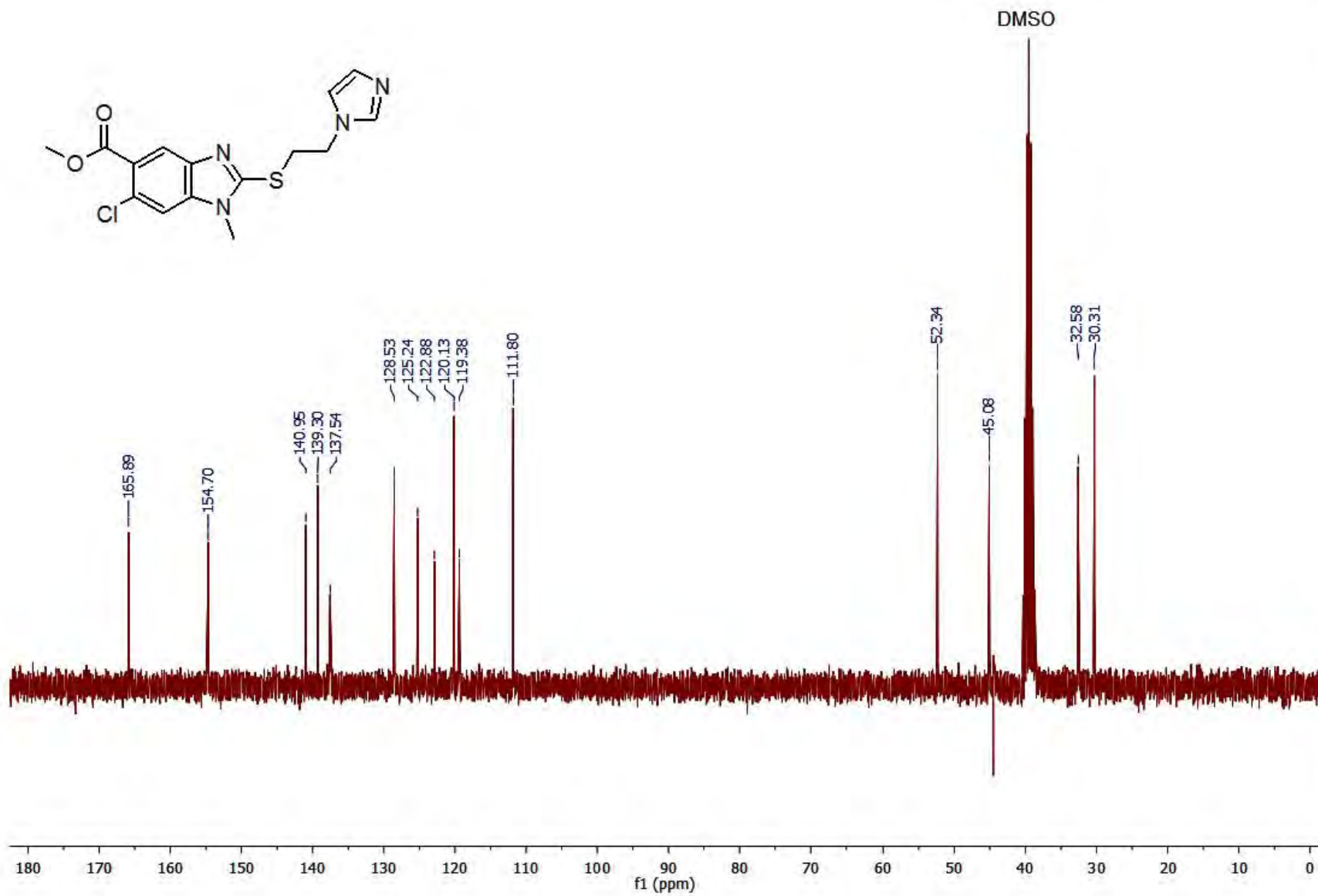
Espectro 35. RMN ¹³C (75.5 MHz) del 2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1-metil-1H-bencimidazol-5-carboxilato de metilo (S2J12)



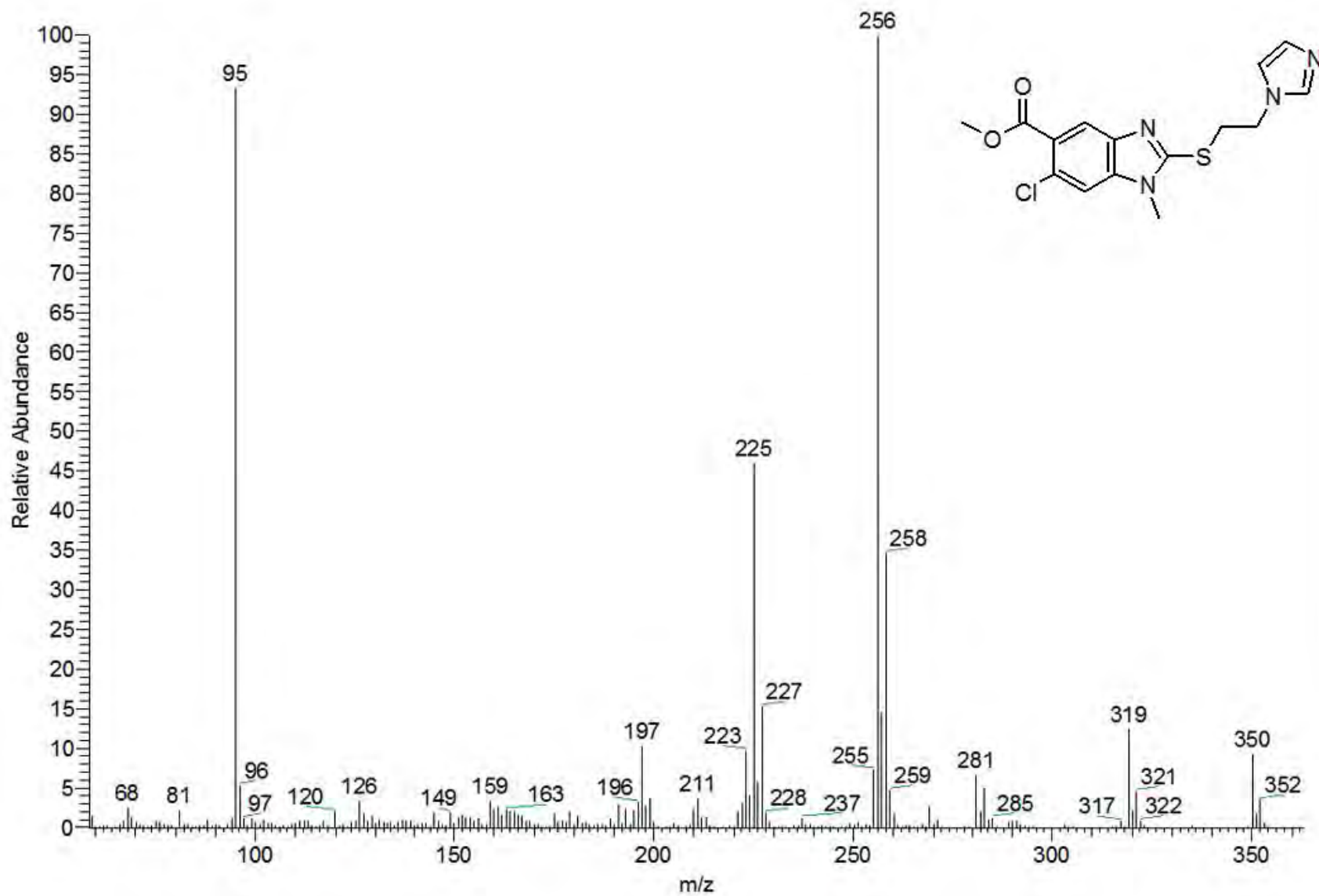
Espectro 36. EM (IE) del 2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-5-carboxilato de metilo (S2J12)



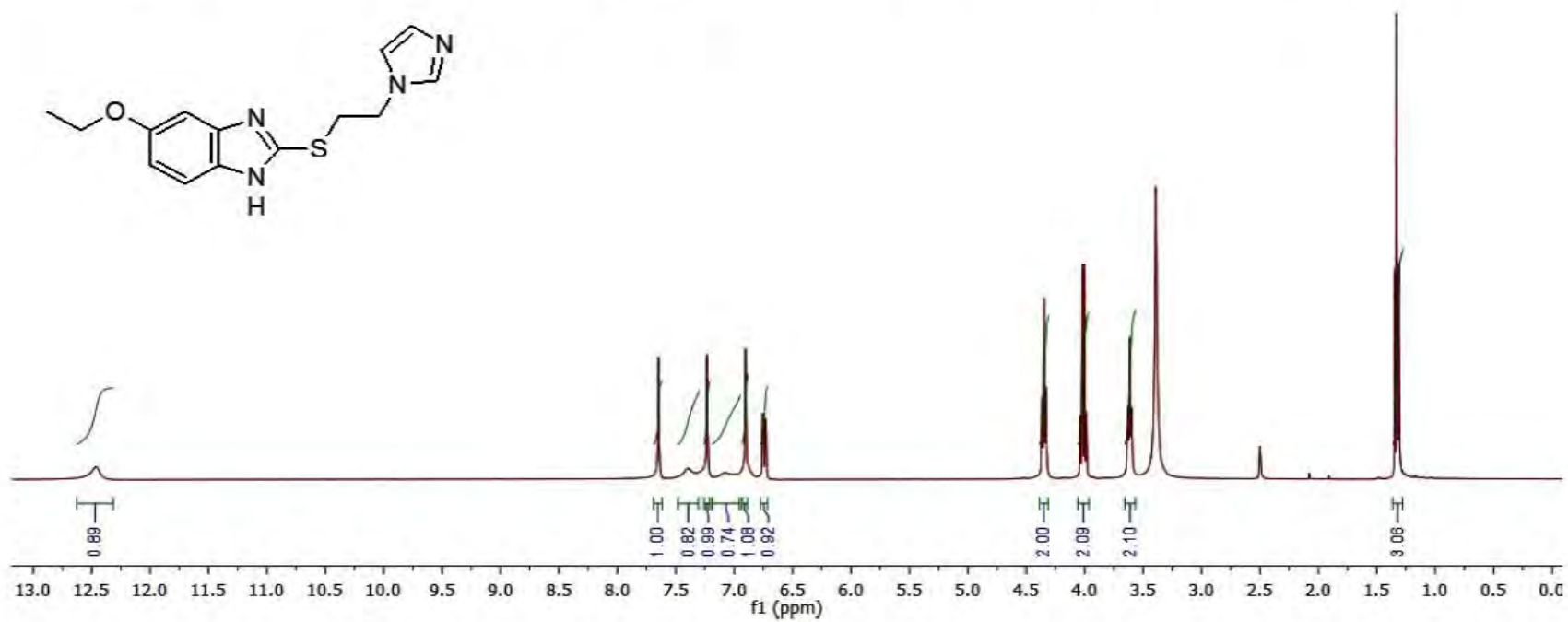
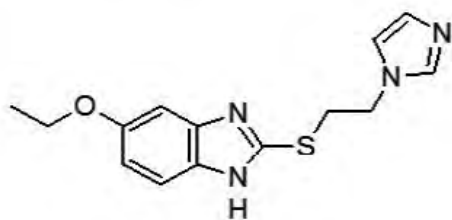
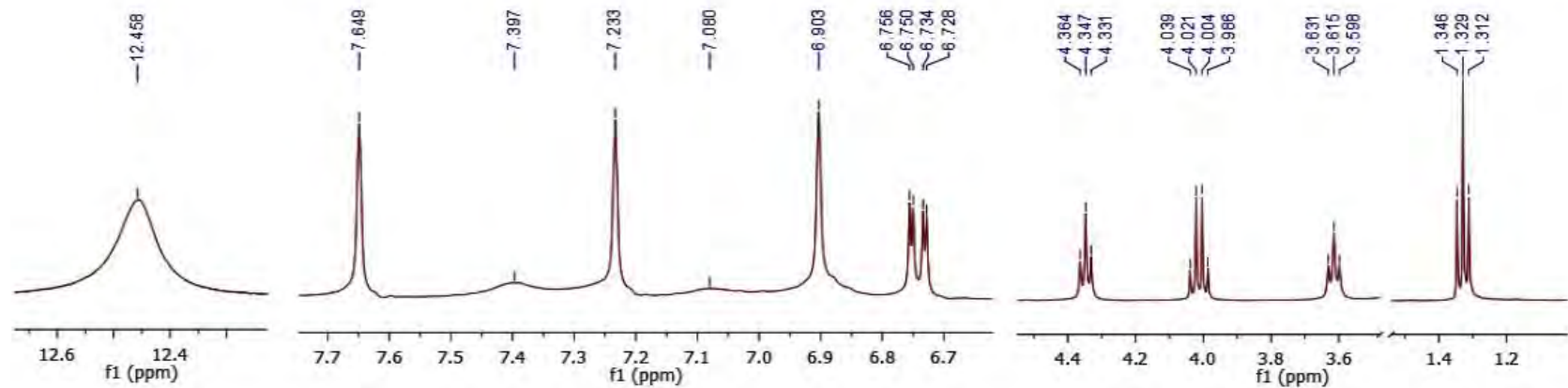
Espectro 37. RMN ^1H (300 MHz) del 6-cloro-2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1-metil-1*H*-bencimidazol-5-carboxilato de metilo (S2J13)



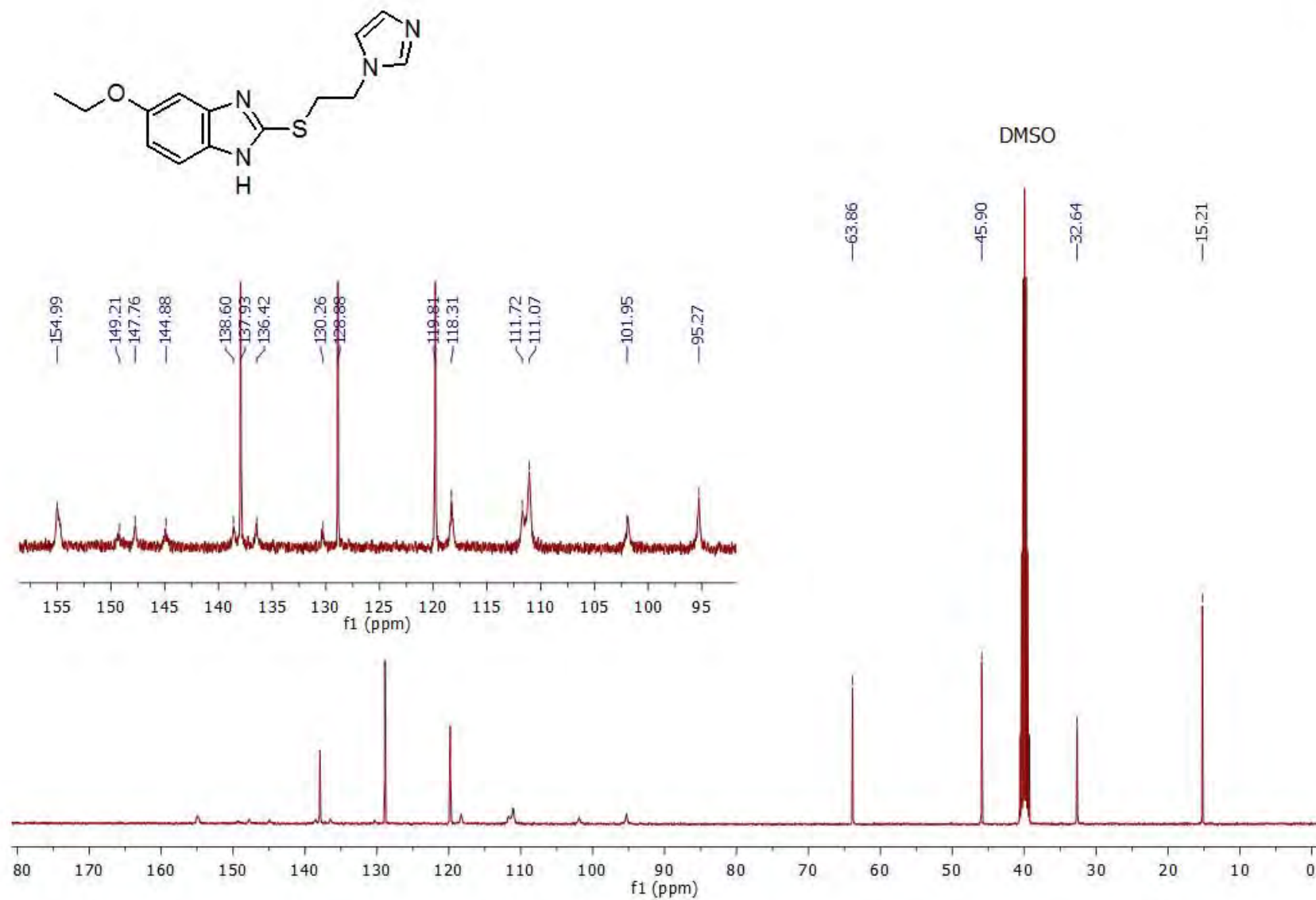
Espectro 38. RMN ¹³C (75.5 MHz) del 6-cloro-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol-5-carboxilato de metilo (S2J13)

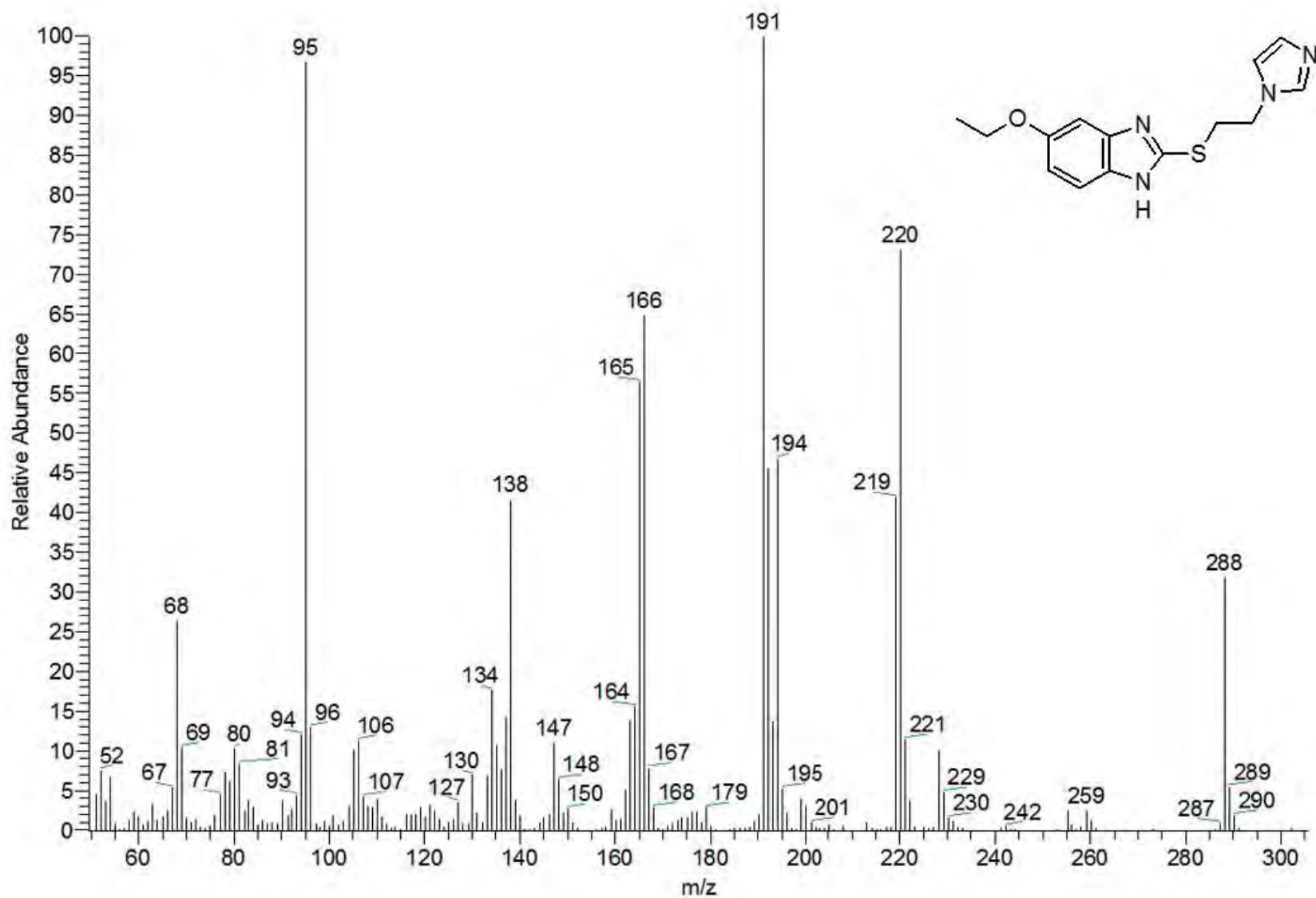


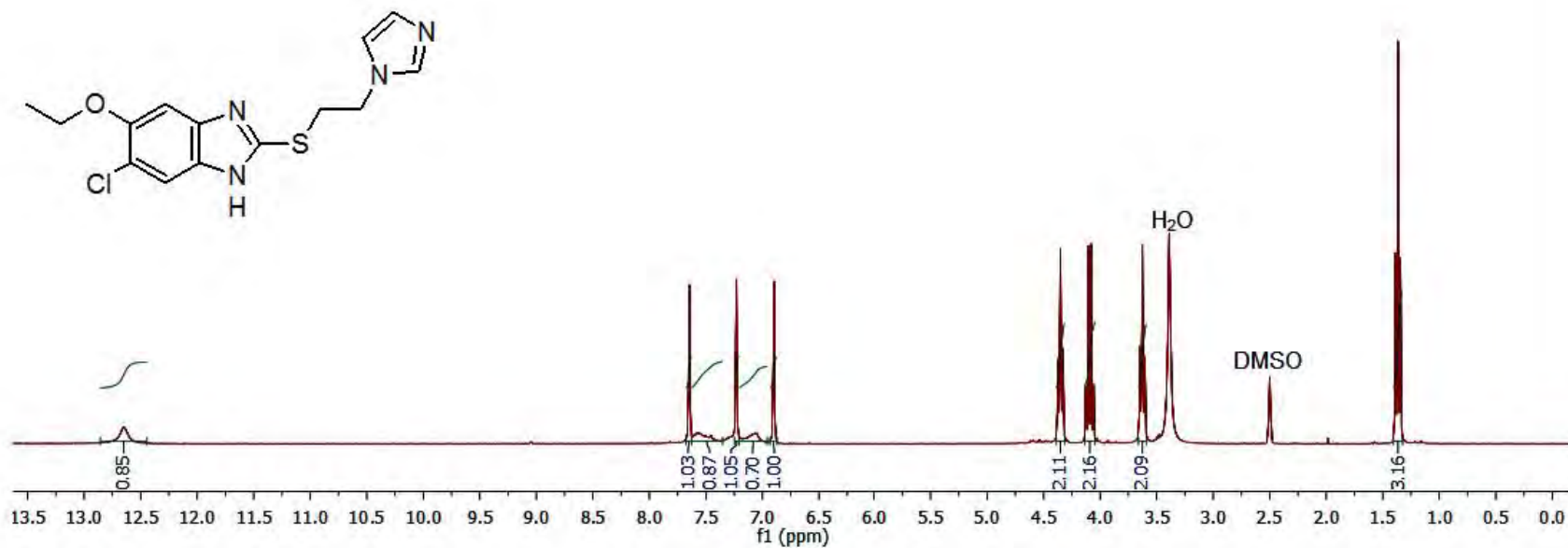
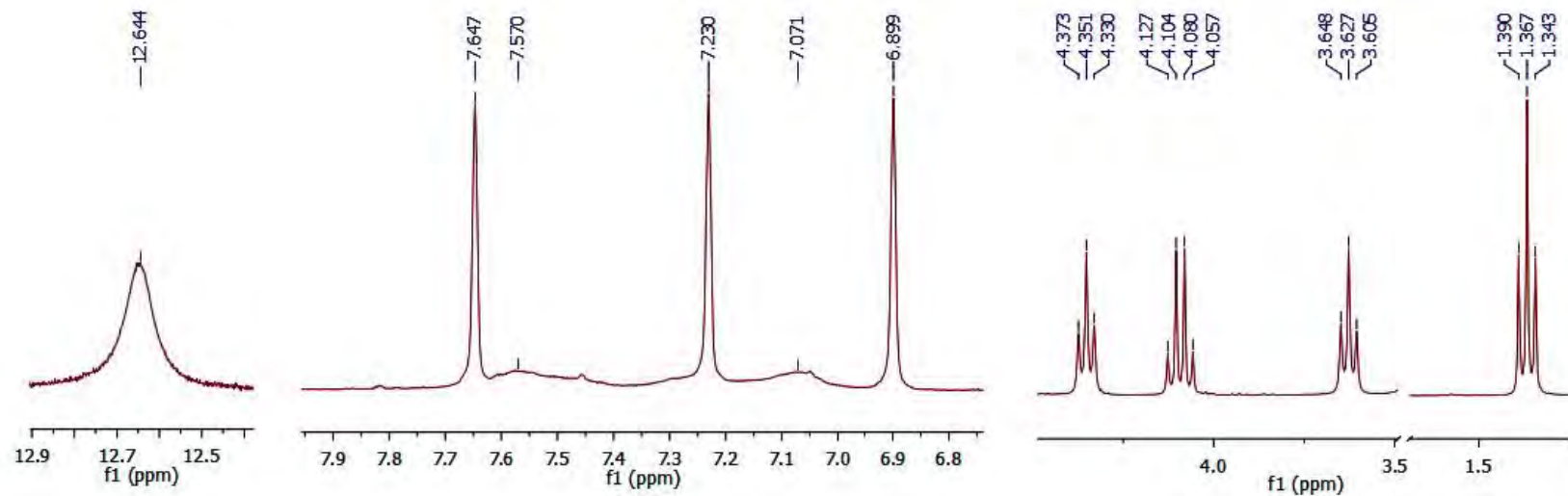
Espectro 39. EM (IE) del 6-cloro-2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1-metil-1H-bencimidazol-5-carboxilato de metilo (S2J13)



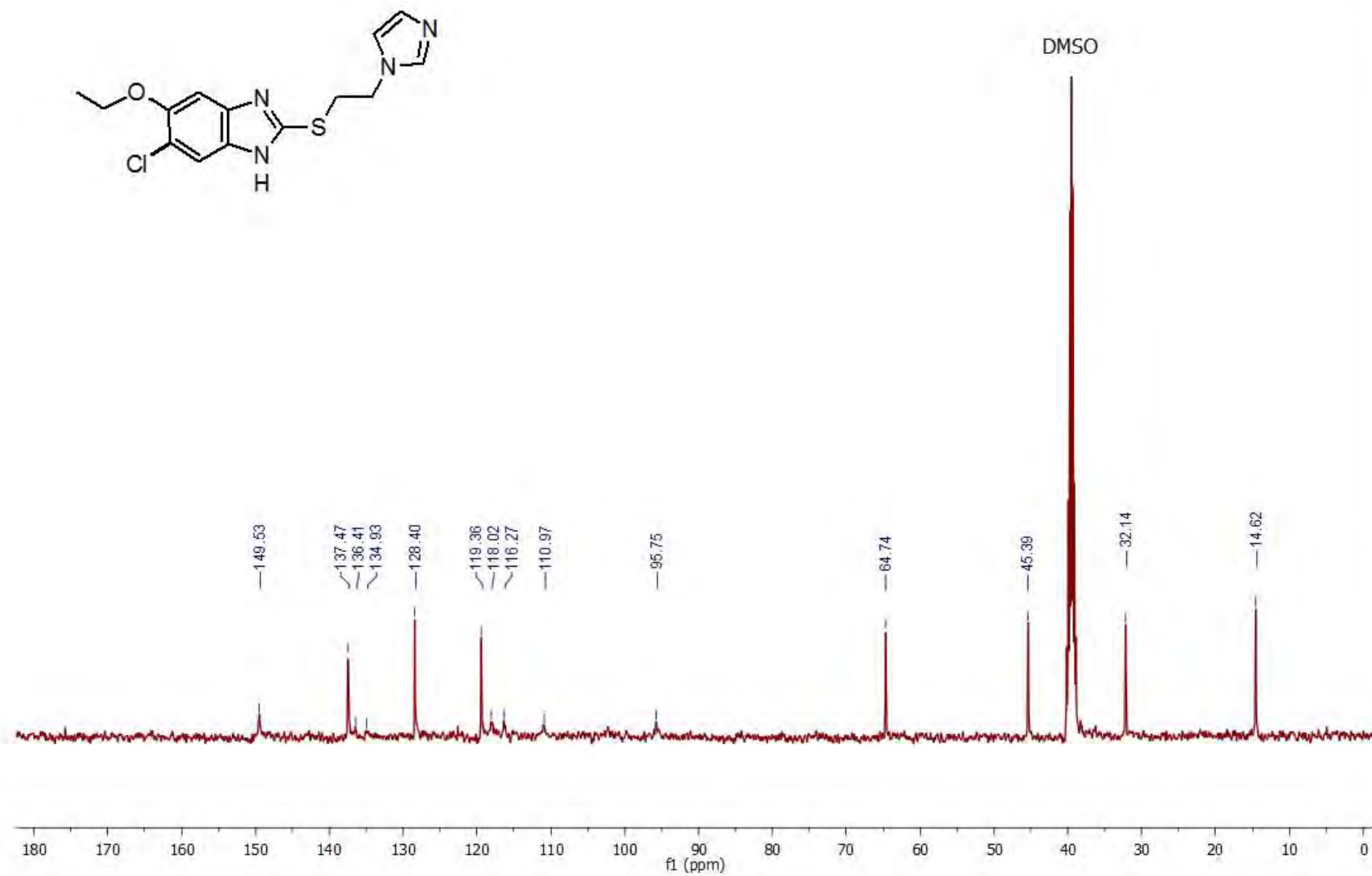
Espectro 40. RMN ^1H (400 MHz) del 5-etoxi-2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1*H*-bencimidazol (**S3J14**)

Espectro 41. RMN ¹³C (75.5 MHz) del 5-etoxi-2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1H-benzimidazol (S3J14)

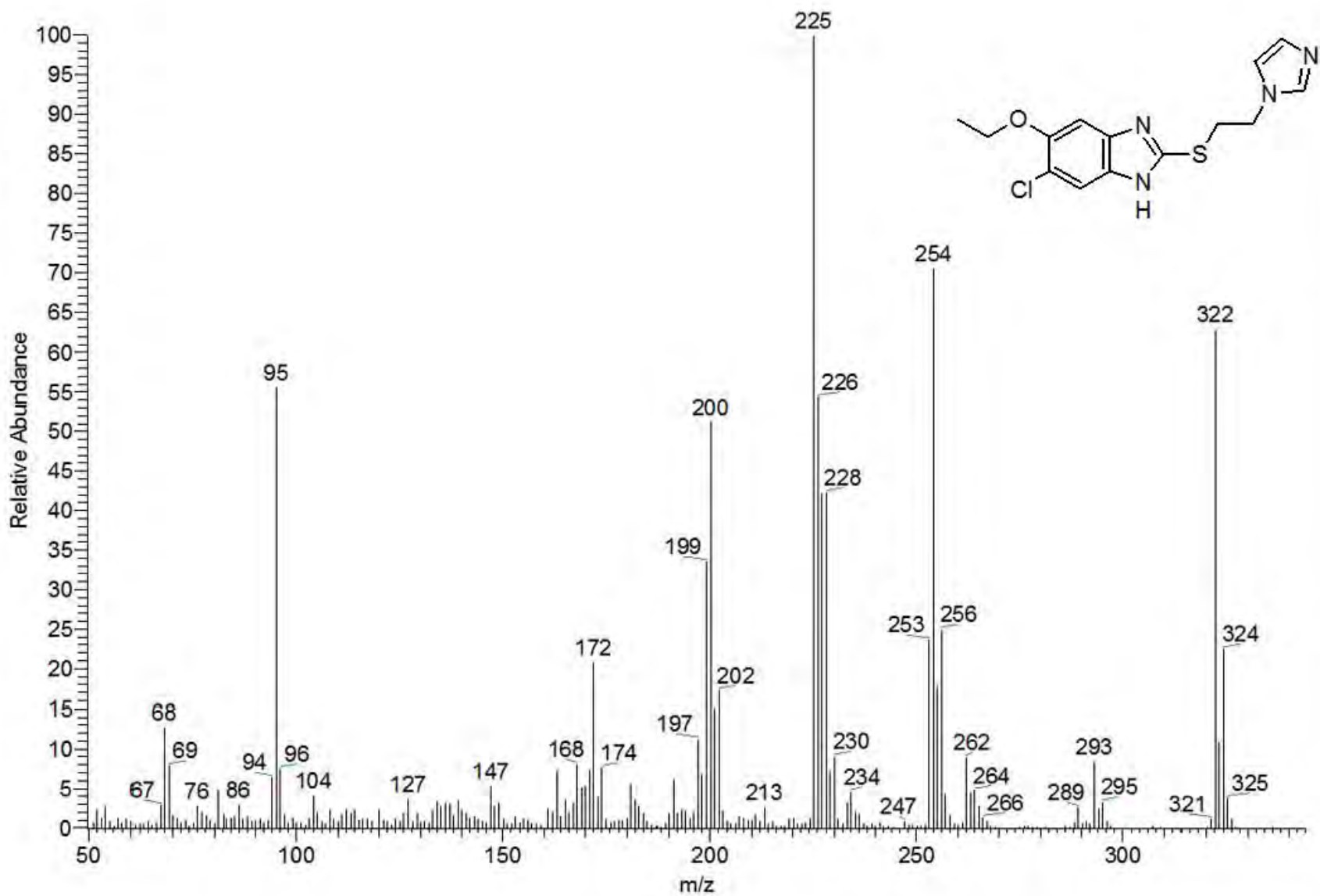
Espectro 42. EM (IE) del 5-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (**S3J14**)

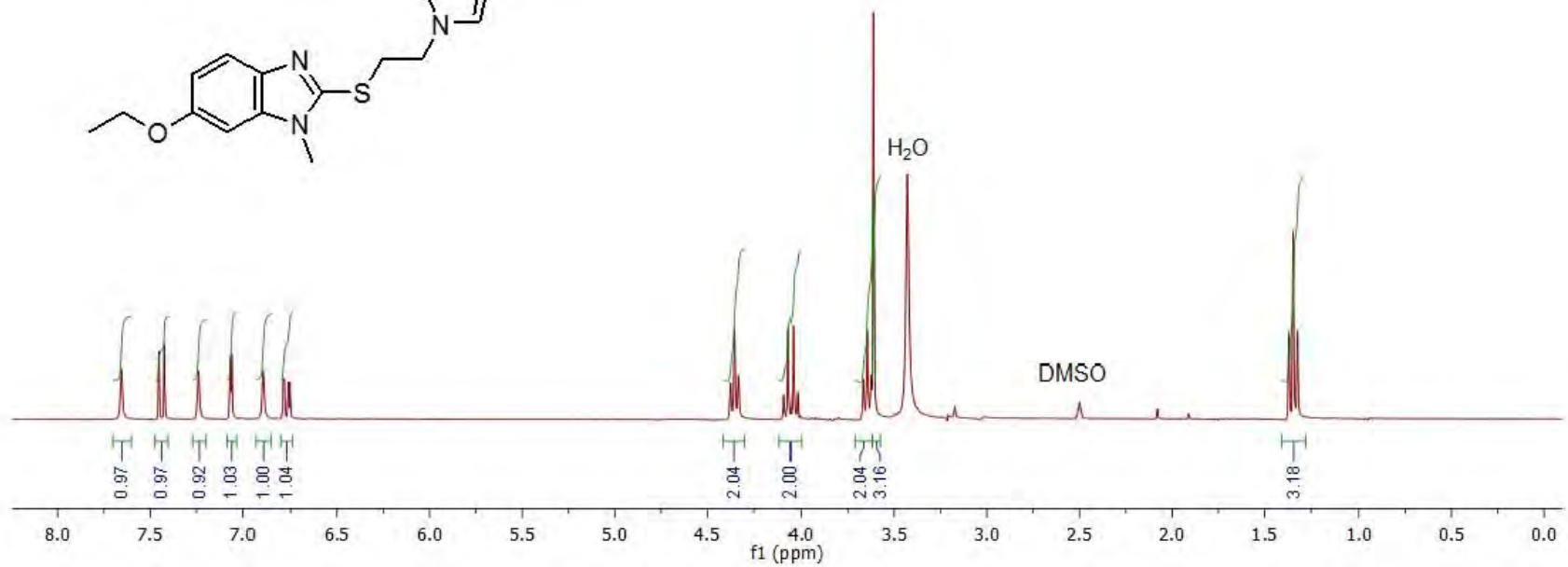
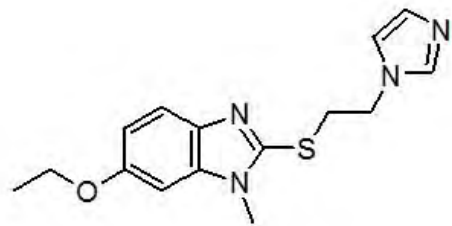
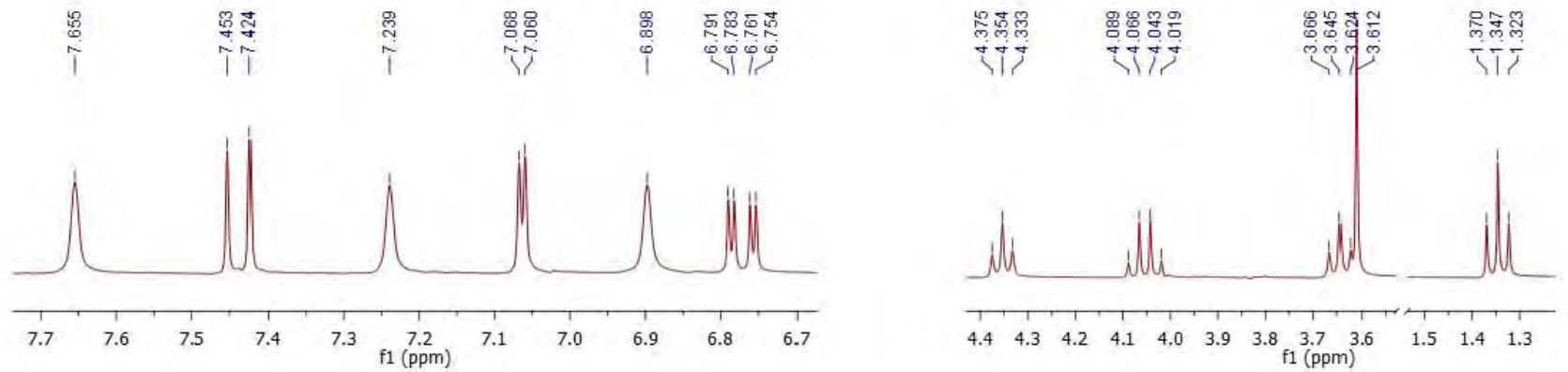


Espectro 43. RMN ^1H (300 MHz) del 6-cloro-5-etoxi-2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1*H*-bencimidazol (**S3J15**)

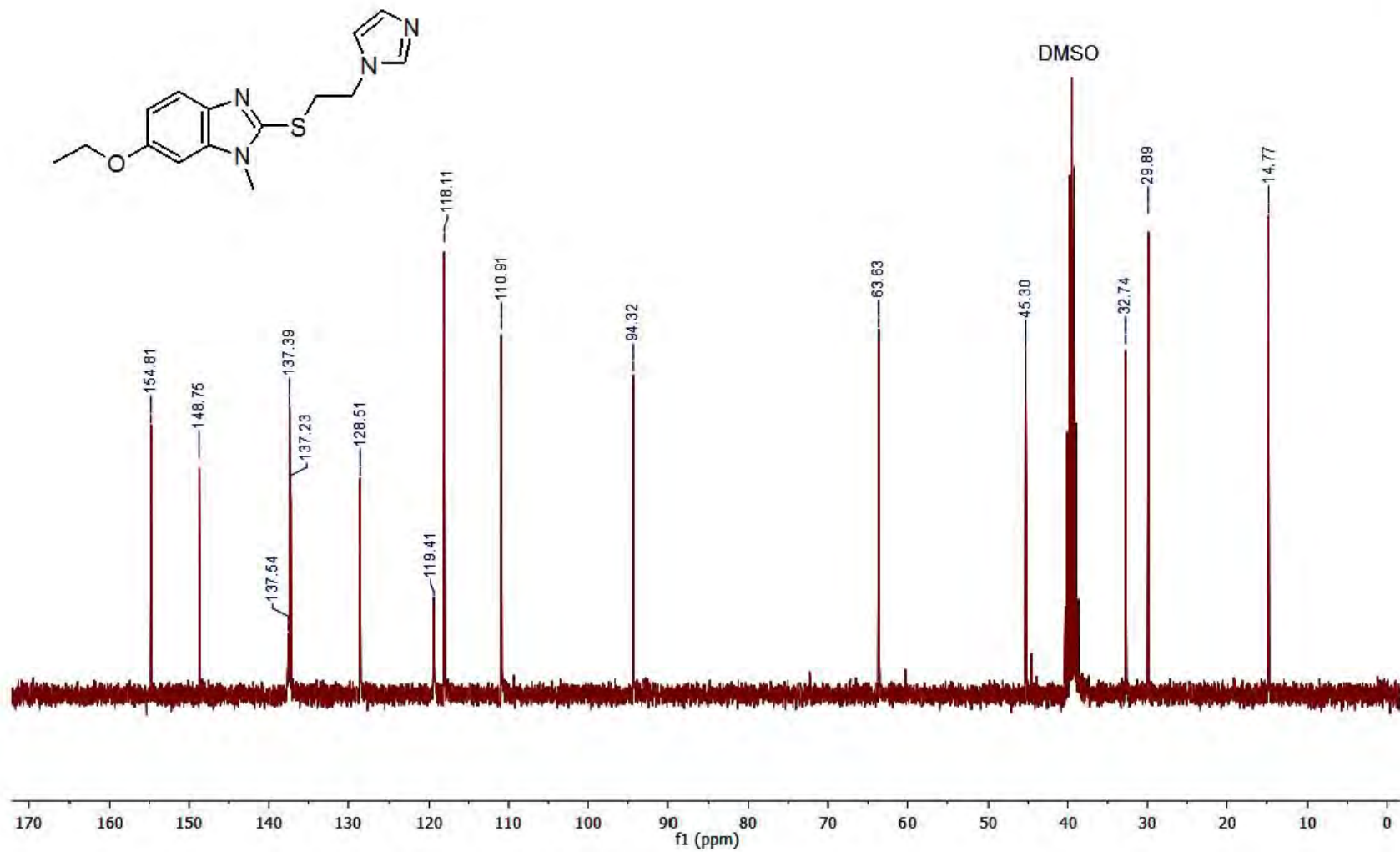
Espectro 44. RMN ¹³C (75.5 MHz) del 6-cloro-5-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-benzimidazol (S3J15)

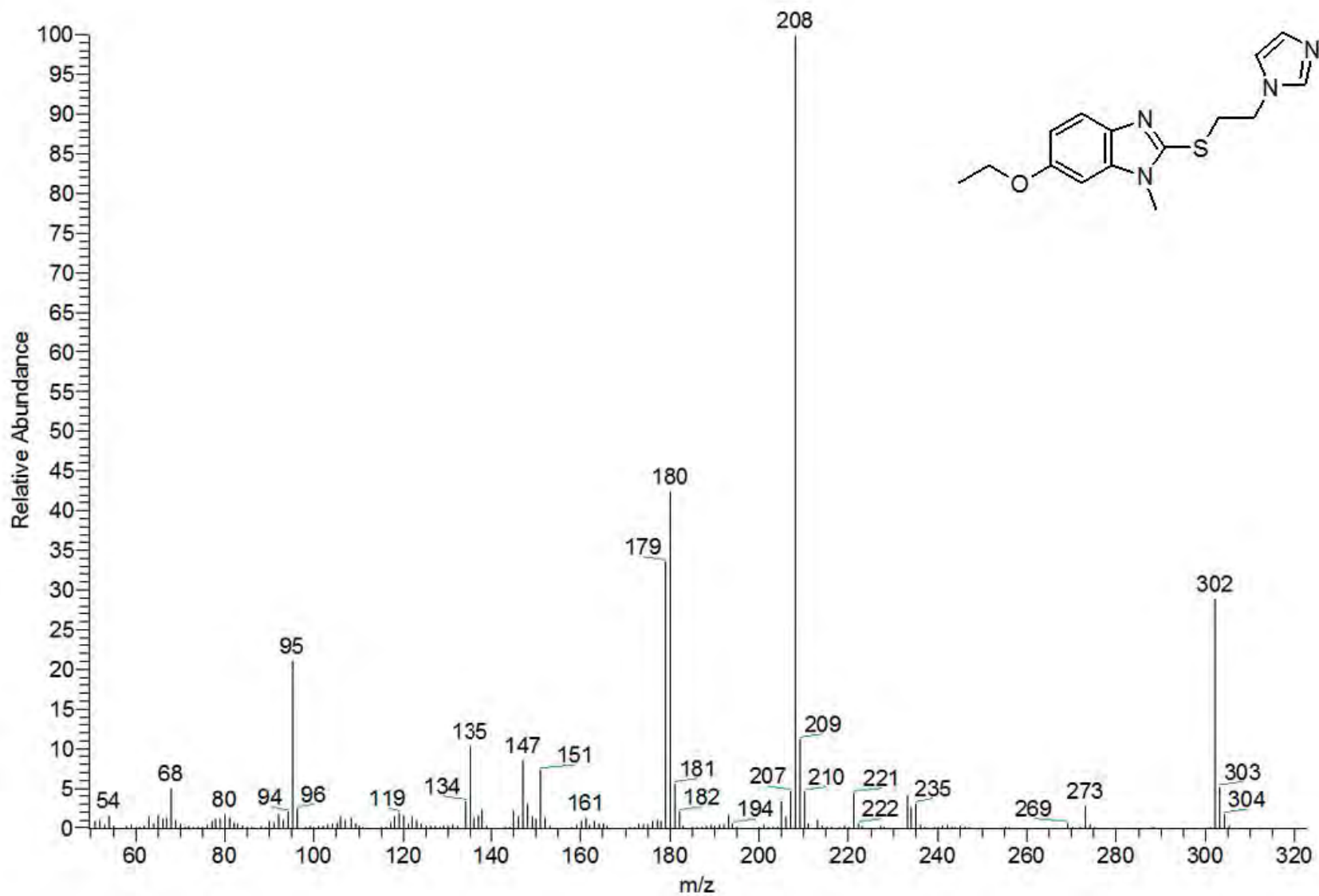
210

Espectro 45. EM (IE) del 6-cloro-5-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1H-bencimidazol (**S3J15**)

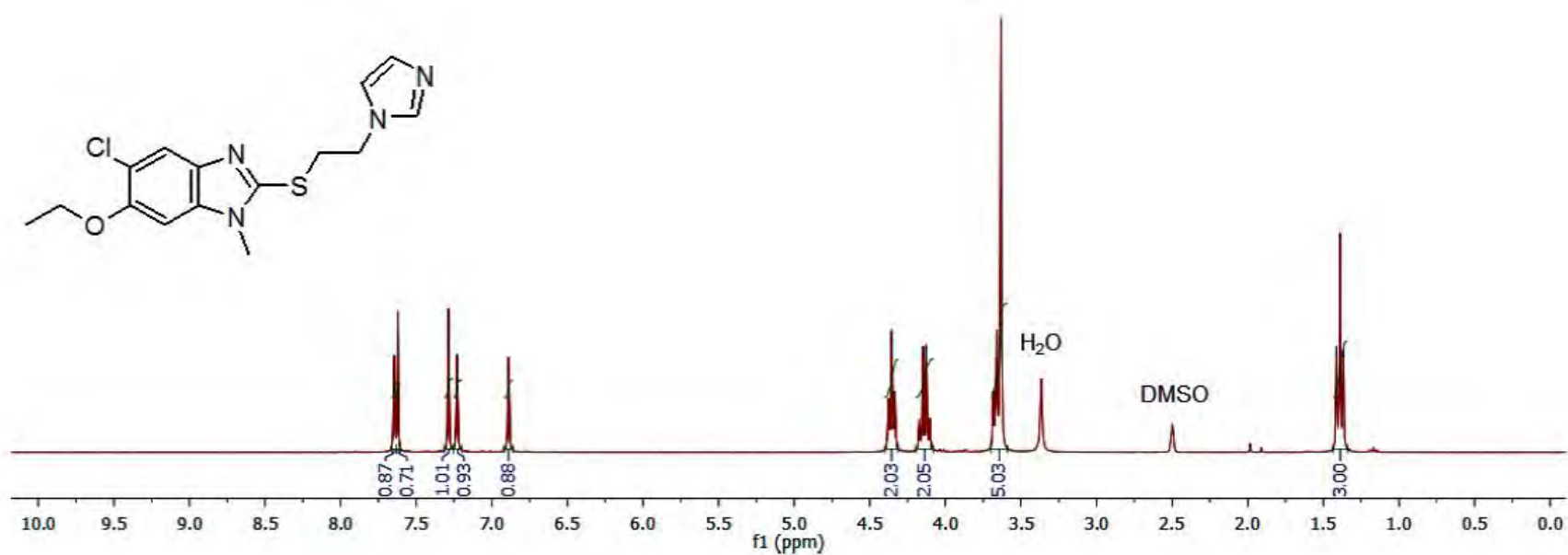
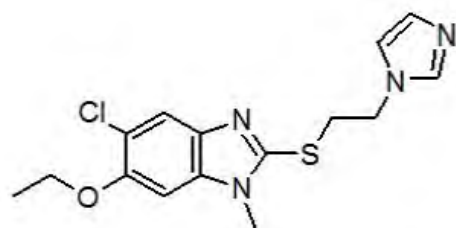
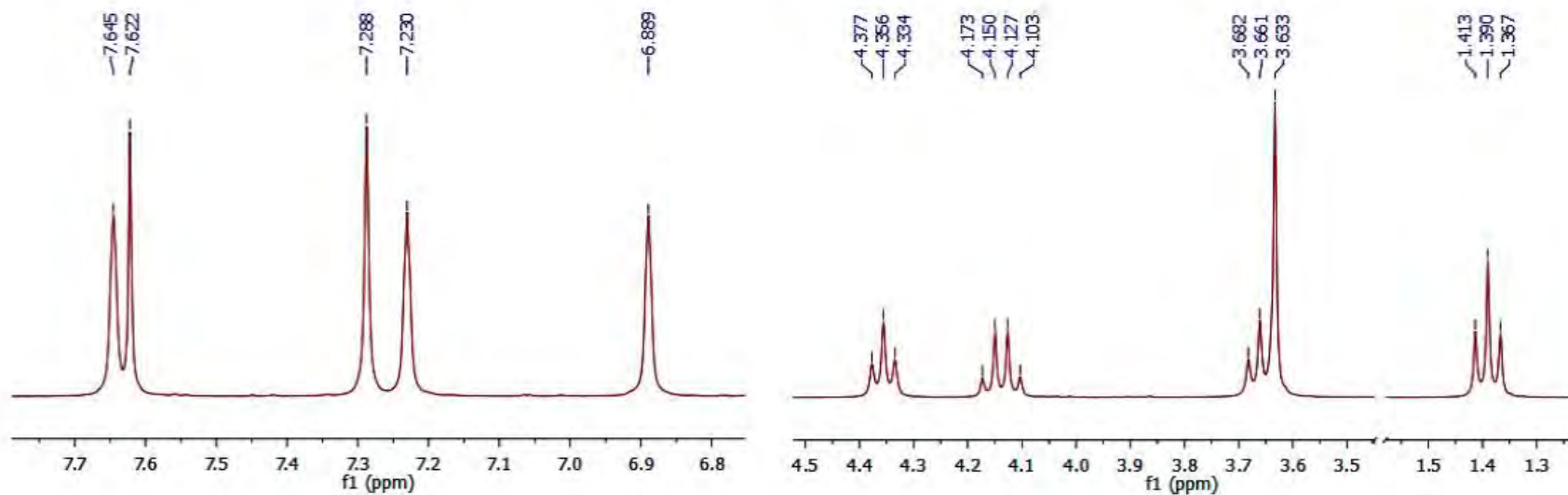


Espectro 46. RMN ^1H (300 MHz) del 6-etoxi-2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1-metil-1*H*-bencimidazol (**S3J16**)

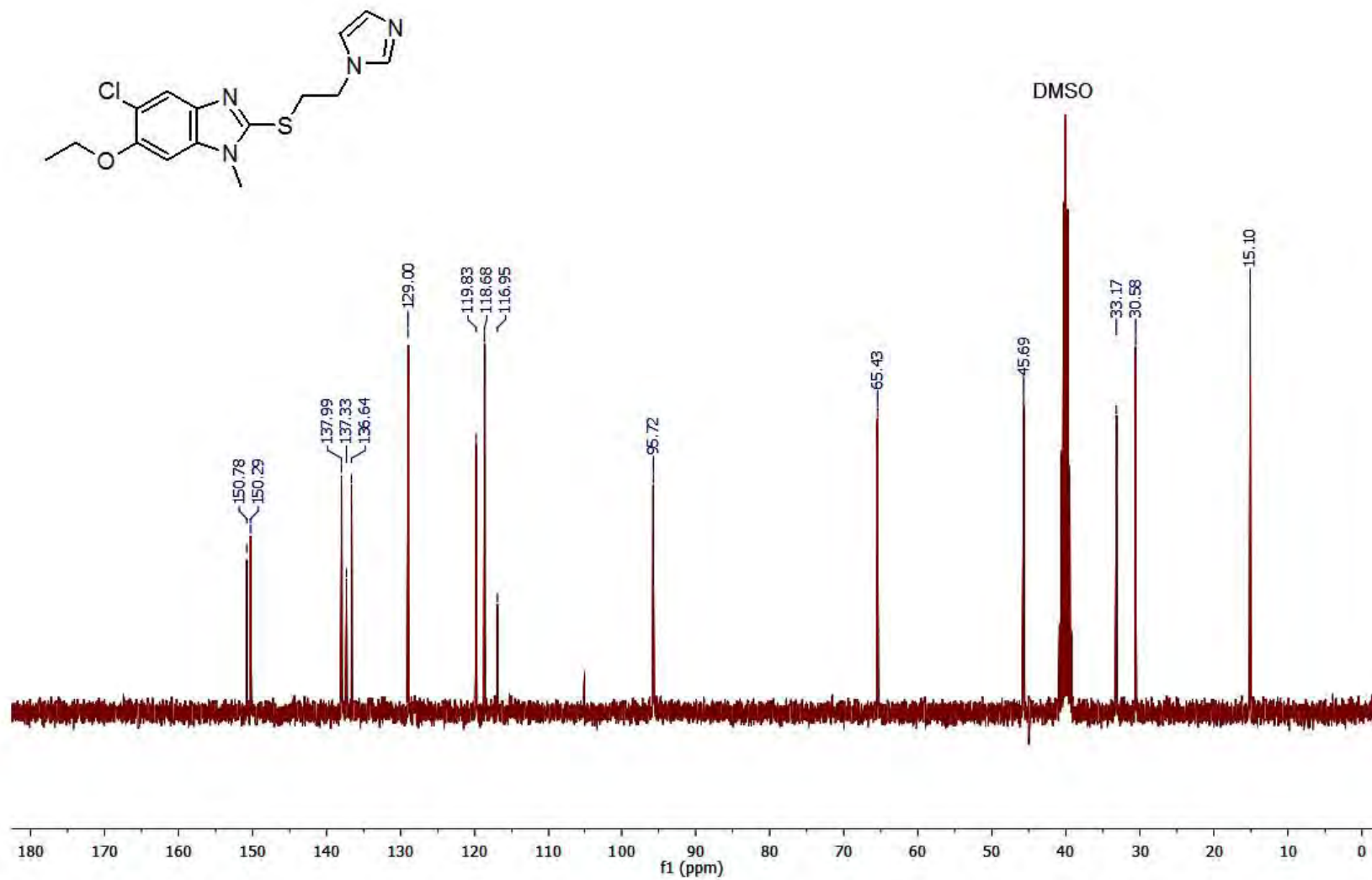
Espectro 47. RMN ¹³C (75.5 MHz) del 6-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S3J16)



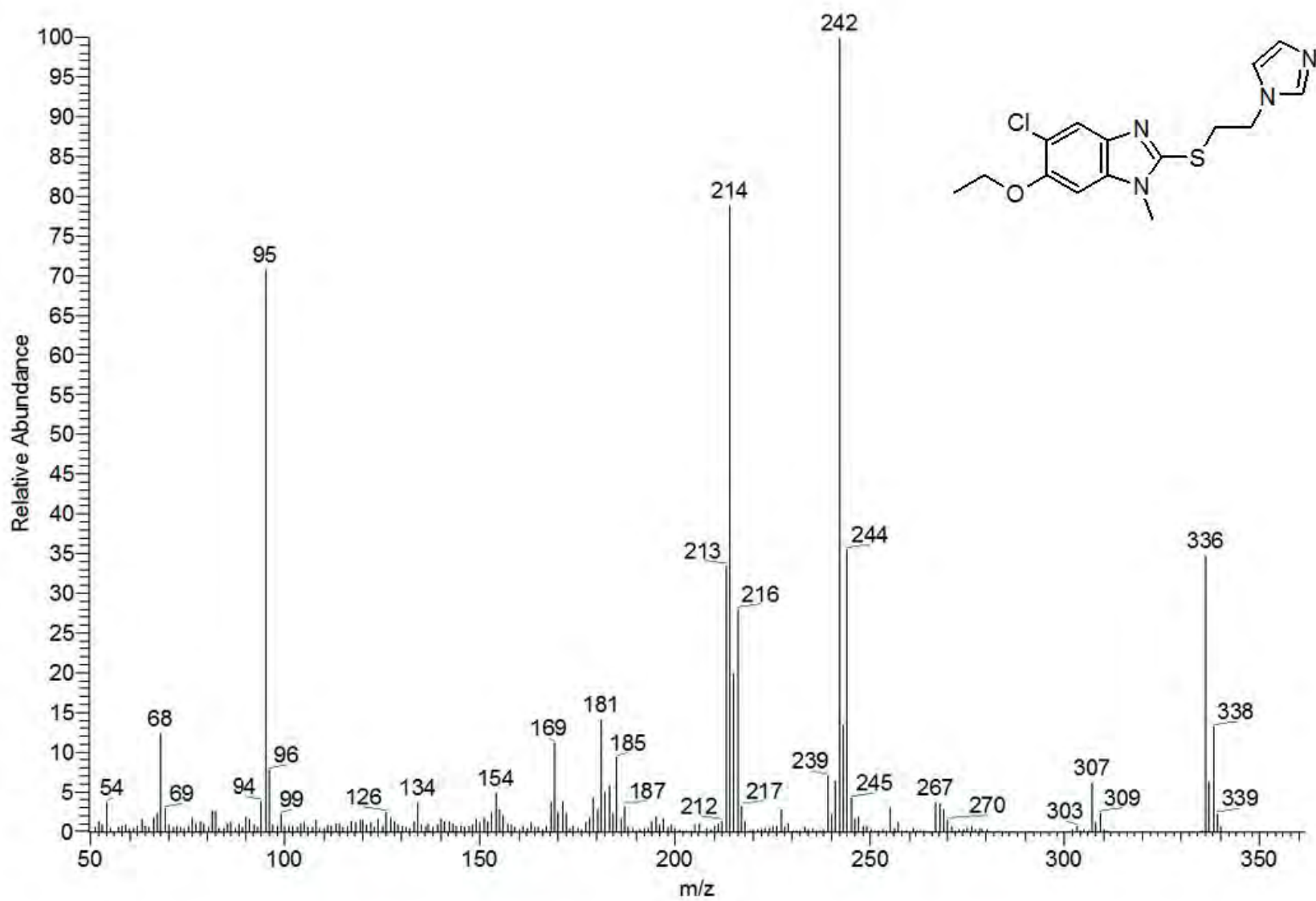
Espectro 48. EM (IE) del 6-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S3J16)

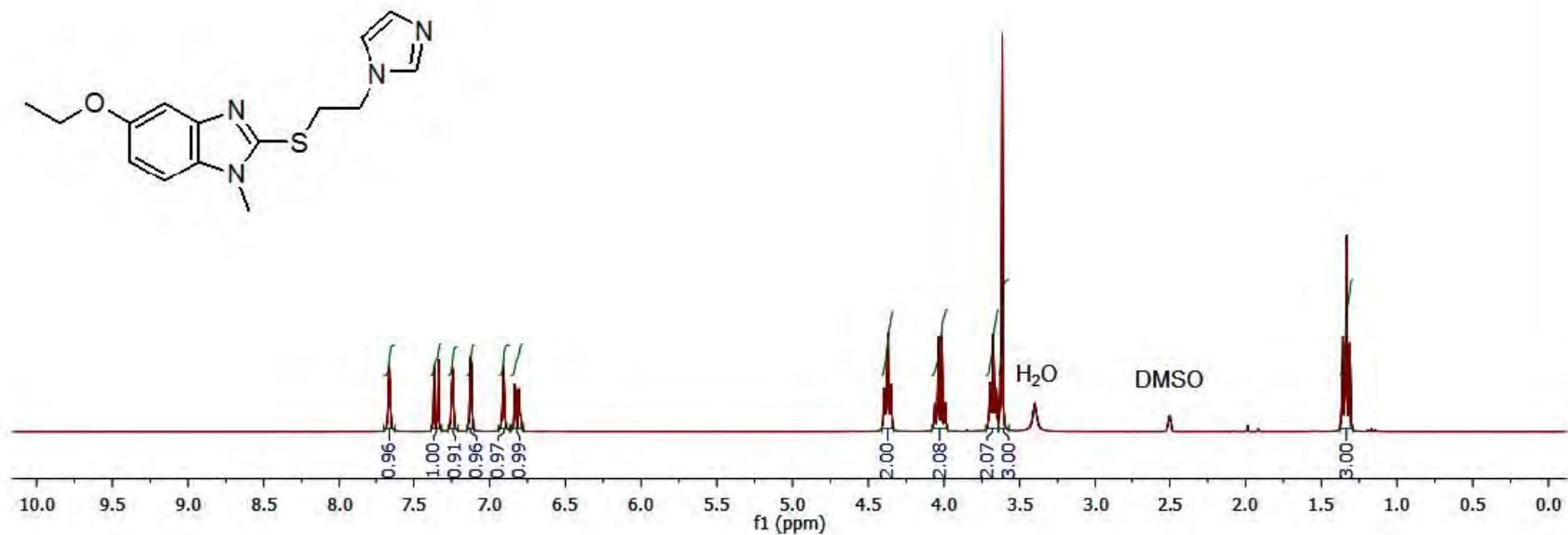
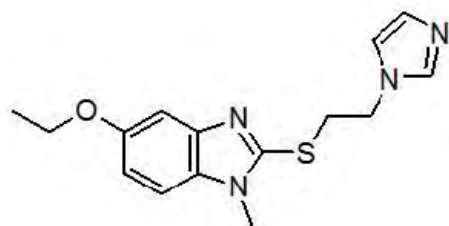
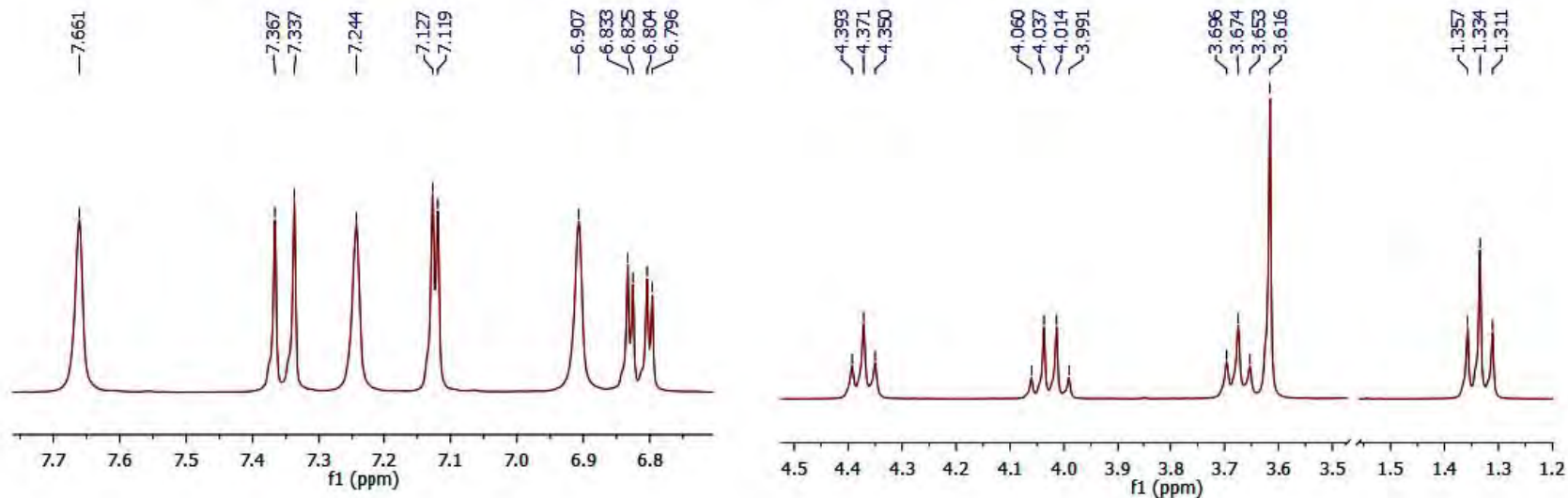


Espectro 49. RMN ^1H (300 MHz) del 5-cloro-6-etoxi-2-[(2-(1*H*-imidazol-1-il)etil)sulfanil]-1-metil-1*H*-bencimidazol (**S3J17**)

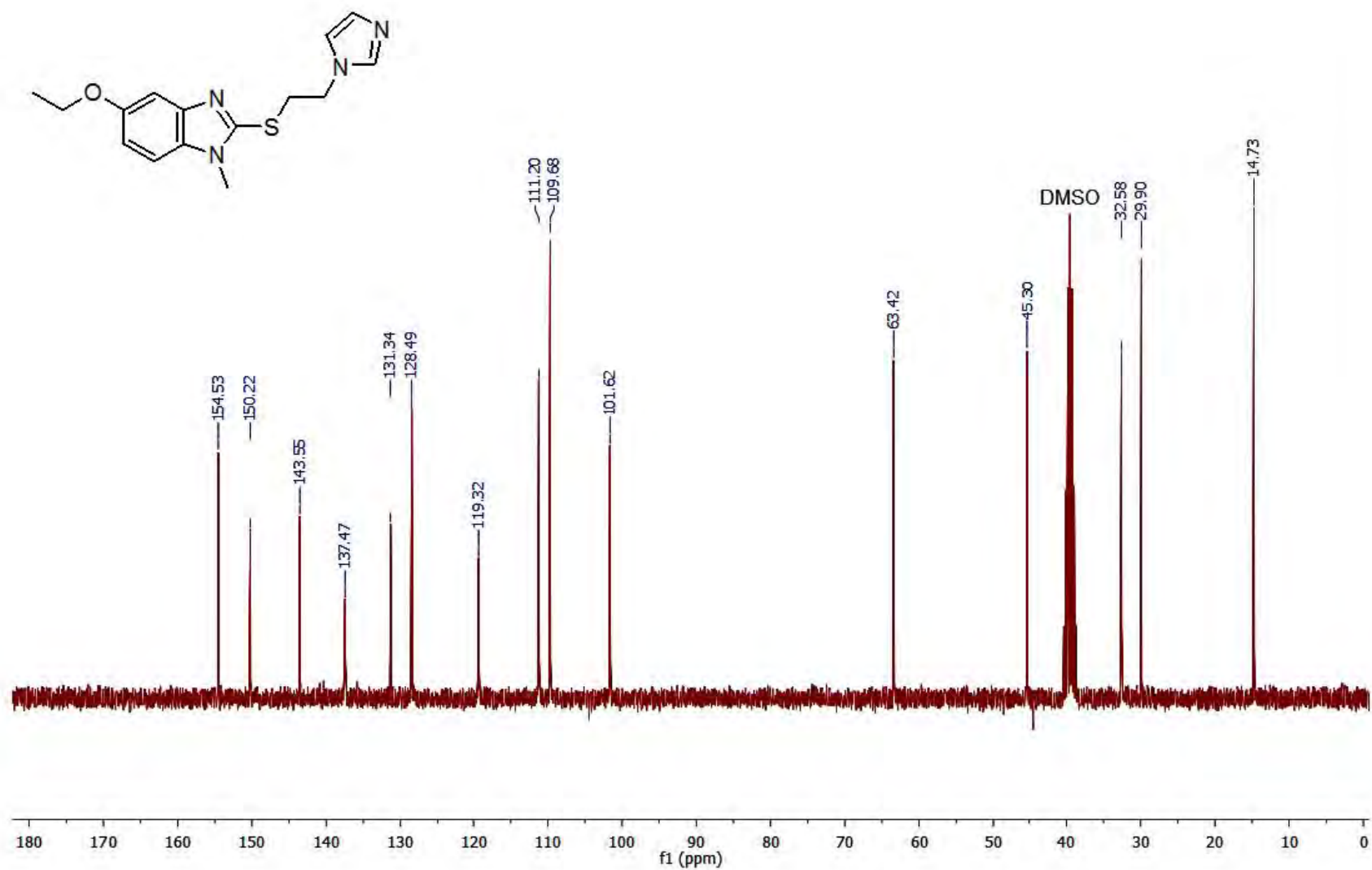


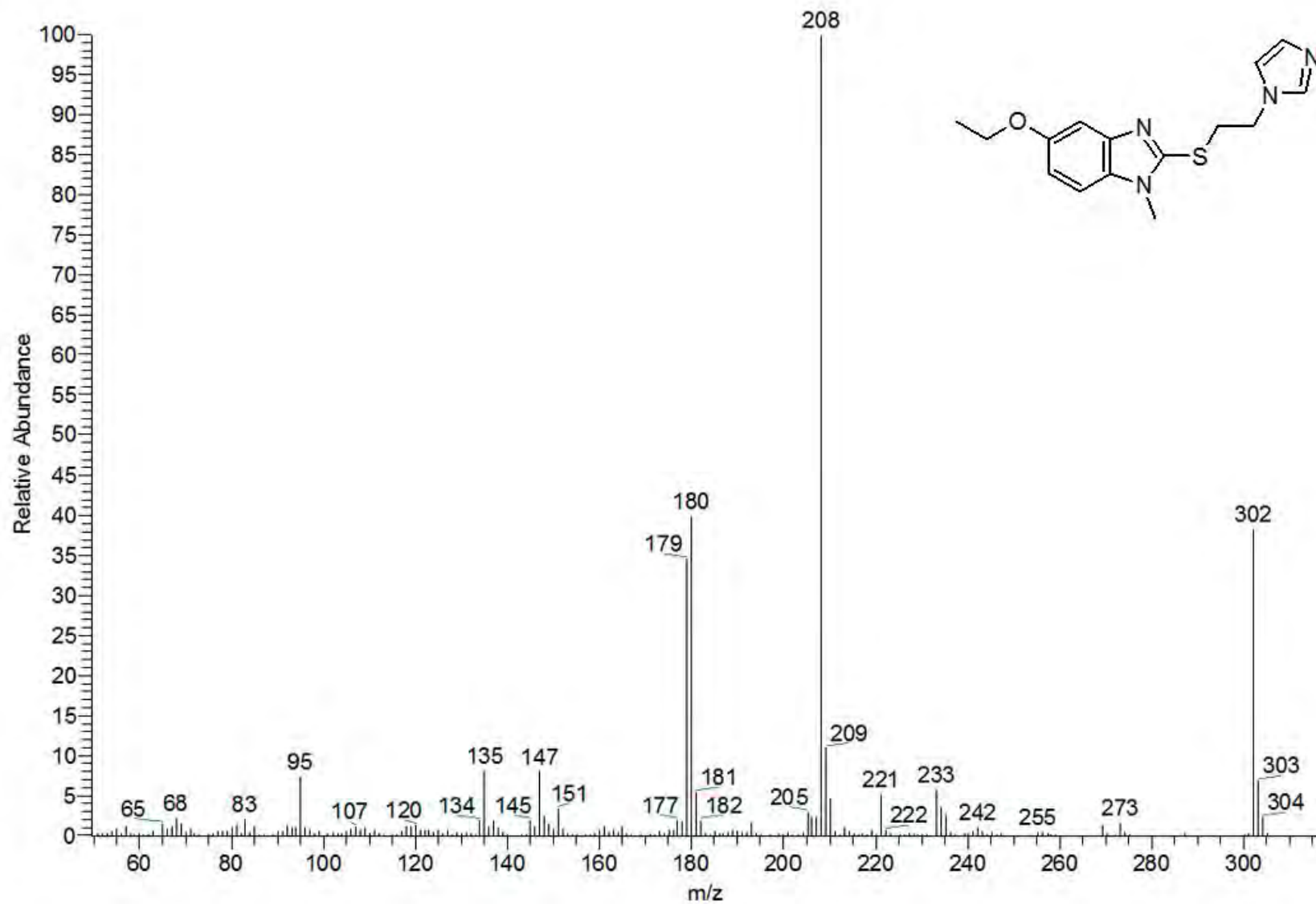
Espectro 50. RMN ¹³C (75.5 MHz) del 5-cloro-6-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (**S3J17**)

Espectro 51. EM (IE) del 5-cloro-6-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (**S3J17**)

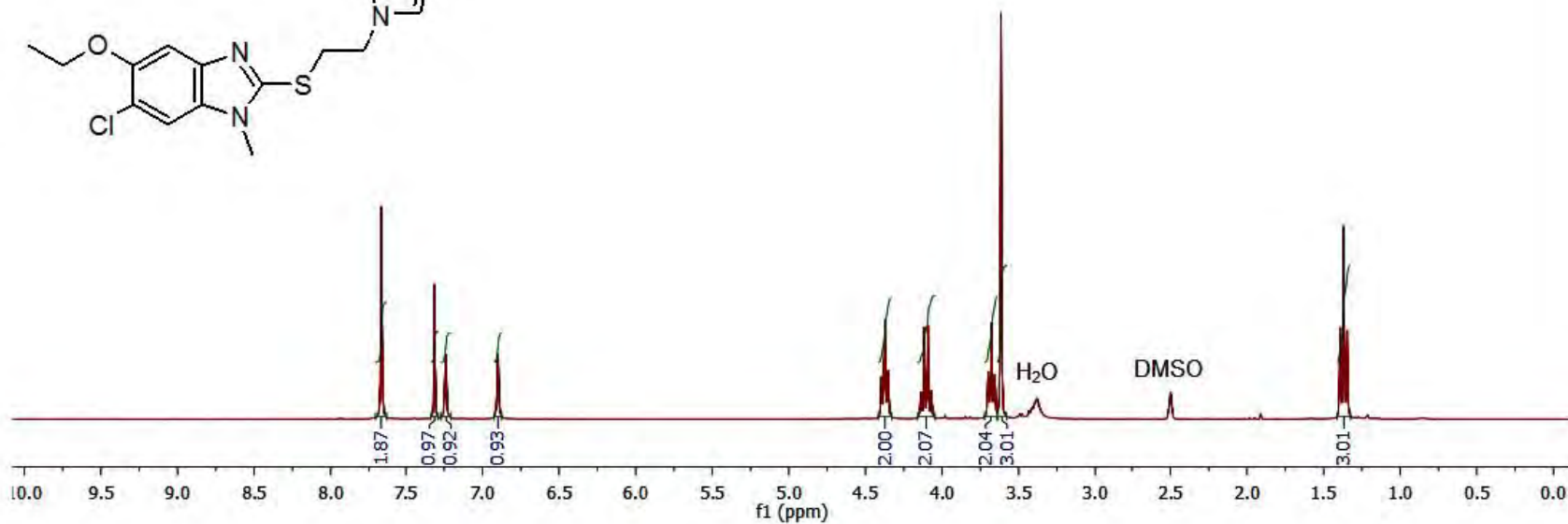
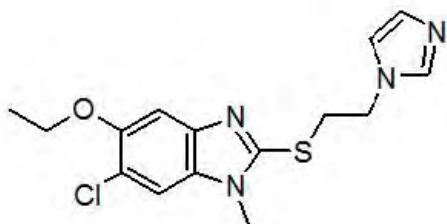
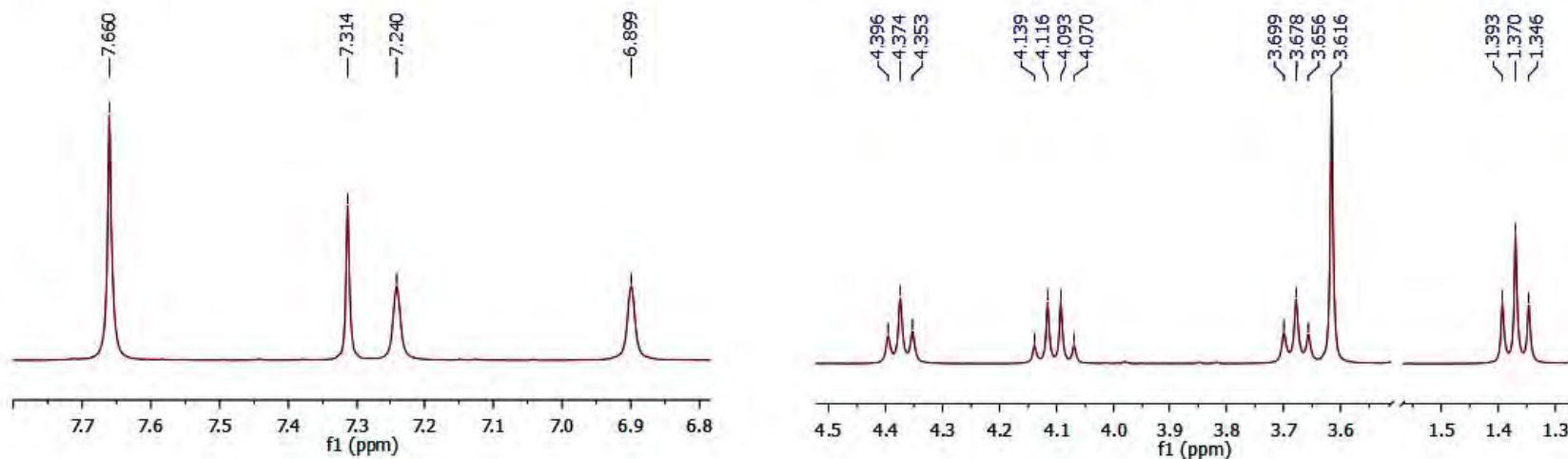


Espectro 52. RMN ¹H (300 MHz) del 5-etoxi-2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1-metil-1*H*-bencimidazol (**S3J18**)

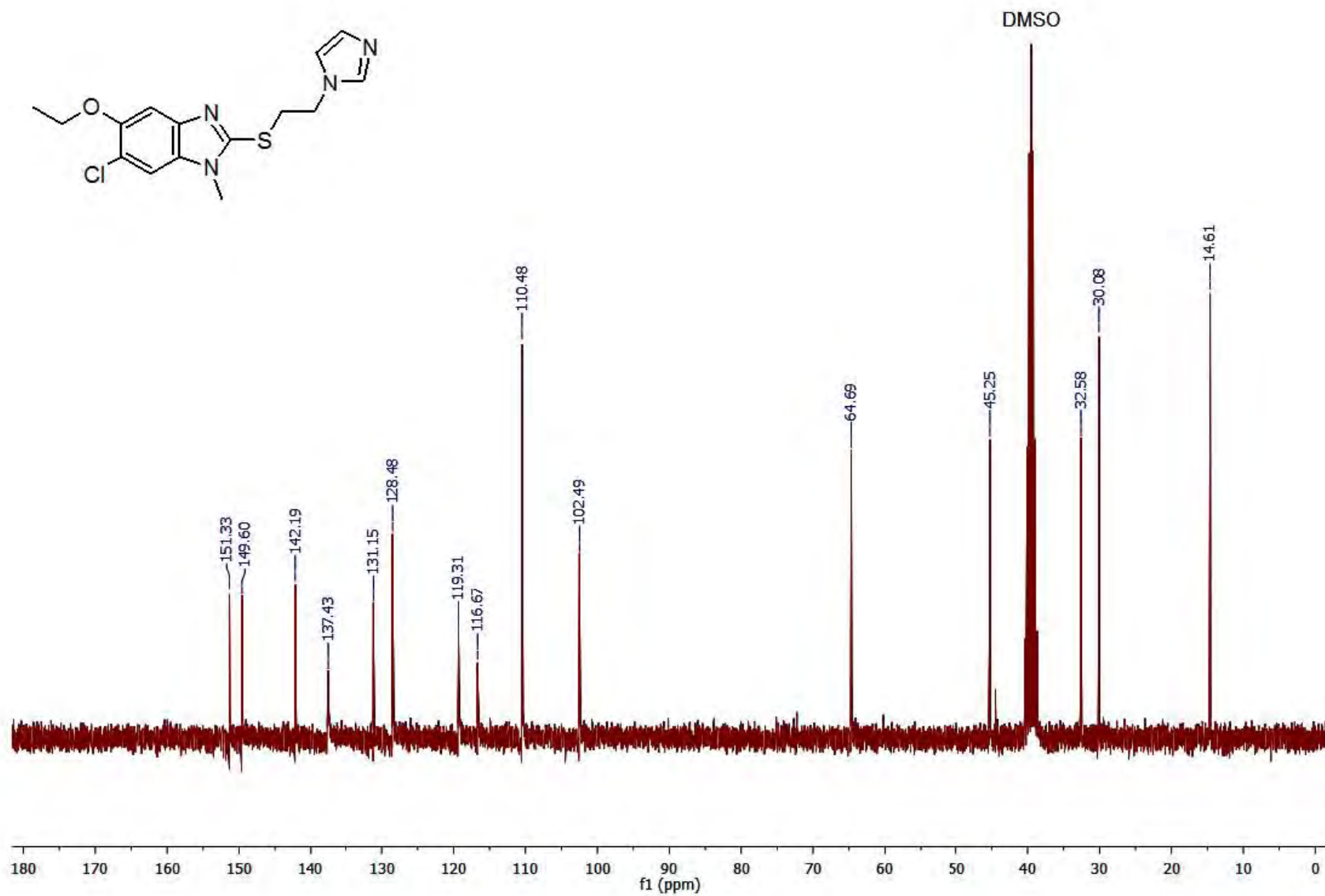
Espectro 53. RMN ¹³C (75.5 MHz) del 5-etoxi-2-[[2-(1H-imidazol-1-yl)etil]sulfanil]-1-metil-1H-bencimidazol (S3J18)



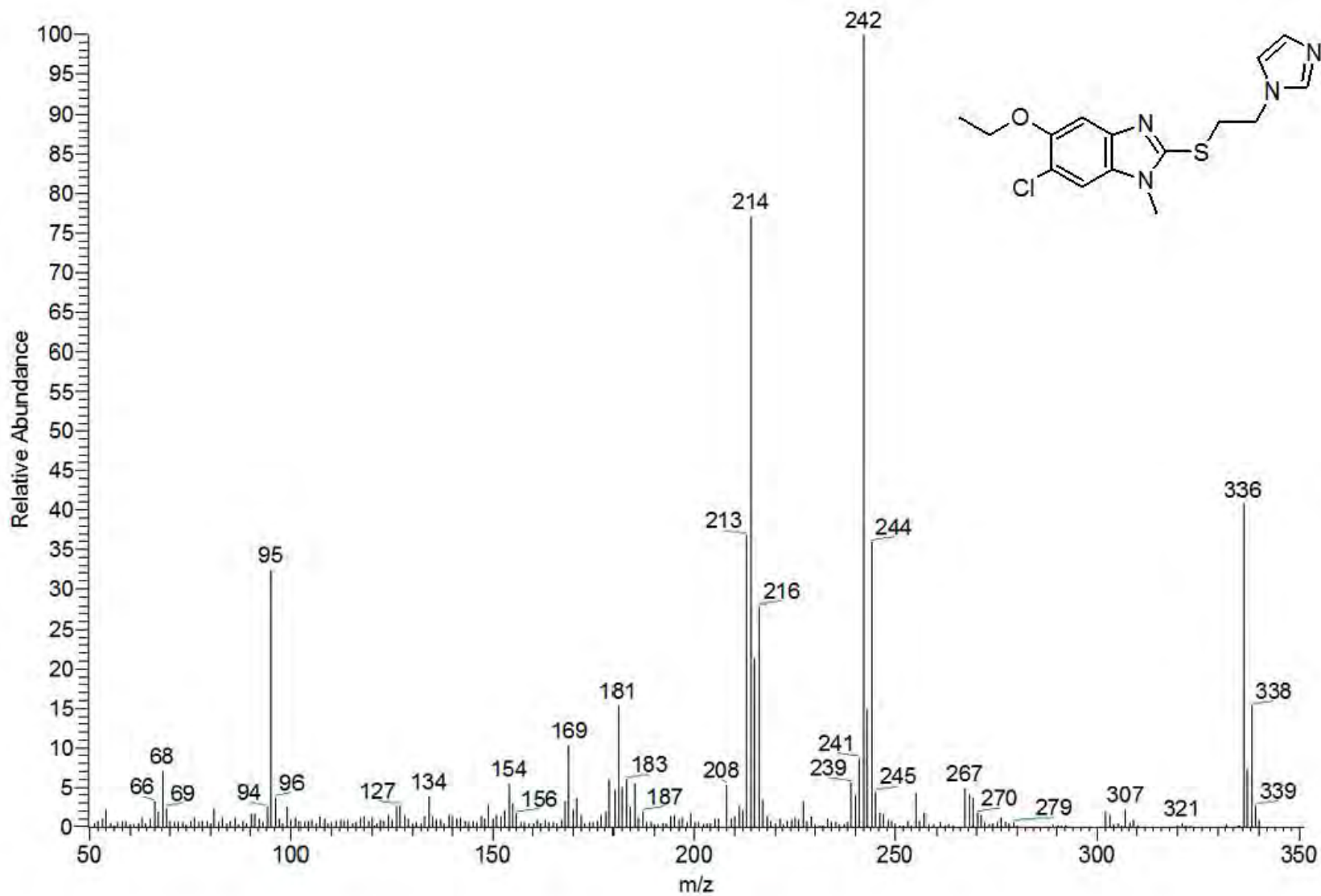
Espectro 54. EM (IE) del 5-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (S3J18)



Espectro 55. RMN ^1H (300 MHz) del 6-cloro-5-etoxi-2-[[2-(1*H*-imidazol-1-il)etil]sulfanil]-1-metil-1*H*-bencimidazol (**S3J19**)



Espectro 56. RMN ¹³C (75.5 MHz) del 6-cloro-5-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (**S3J19**)

Espectro 57. EM (IE) del 6-cloro-5-etoxi-2-[[2-(1H-imidazol-1-il)etil]sulfanil]-1-metil-1H-bencimidazol (**S3J19**)

PUBLICACIONES



Towards a systematic characterization of the antiprotozoal activity landscape of benzimidazole derivatives

Jaime Pérez-Villanueva^a, Radleigh Santos^b, Alicia Hernández-Campos^a, Marc A. Giulianotti^b, Rafael Castillo^a, Jose L. Medina-Franco^{b,*}

^aFacultad de Química, Departamento de Farmacia, Universidad Nacional Autónoma de México, México DF 04510, Mexico

^bTorrey Pines Institute for Molecular Studies, 11350 SW Village Parkway, Port St. Lucie, FL 34987, USA

ARTICLE INFO

Article history:

Received 9 August 2010

Revised 31 August 2010

Accepted 6 September 2010

Available online 15 September 2010

Keywords:

Activity cliffs

Giardia

Molecular similarity

SAS maps

Structure–activity relationships

Trichomonas

ABSTRACT

Parasitic infections caused by the protozoa *Trichomonas vaginalis* and *Giardia intestinalis* still represent a major problem in developing countries. Despite the fact that benzimidazoles are promising compounds with activity against both protozoa, systematic studies to characterize and compare their structure–activity relationships (SAR) are limited. Herein, we report a systematic characterization of the SAR of 32 benzimidazoles with activity against *T. vaginalis* and *G. intestinalis*. The analysis was based on pairwise comparisons of the activity similarity and molecular similarity using different molecular representations. Radial, MACCS keys, TGD and piDAPH3 fingerprints were used to develop consensus models of the landscape. The landscapes contained continuous regions and activity cliffs. Two ‘deep consensus activity cliffs’ and several pairs of compounds in smooth regions of the SAR were identified in the landscape of *T. vaginalis*. In contrast, a number of ‘apparent and shallow cliffs’ were found for *G. intestinalis*. Several compounds active for both parasites showed similar SAR suggesting a common mechanism of action. We also identified pairs of structurally similar molecules with dramatic changes in selectivity. Results suggested that while substitution at position 2 on the benzimidazole moiety plays an important role in increasing the potency against both parasites, substitutions at positions 4–7 could influence selectivity. This study represents a first step towards the systematic characterization of the antiprotozoal activity landscape of benzimidazoles, and has direct implications in the future development of other types of quantitative models. The landscape of larger data sets with other biological endpoints can be analyzed using the general approaches used in this work.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Parasitic diseases are still a major health problem in developing countries. Mucosal infections by protozoa infect more than a billion people every year. Among the most common protozoa infections are giardiasis, caused by *Giardia intestinalis*, and trichomonosis, a genitourinary infection caused by *Trichomonas vaginalis*.^{1–3} As part of ongoing efforts to develop compounds as giardicidal and trichomonocidal agents, several benzimidazole derivatives have been synthesized and tested, leading to the identification of compounds in the low nanomolar range.^{4–7} However, systematic and quantitative studies of the SAR of benzimidazoles as antiprotozoal agents are still limited.

Quantitative characterization of the SAR of small molecules plays a key role in lead optimization. To this end, a number of methods can be employed including quantitative structure–activity relationships (QSAR), rule-based methods, neural networks or

pharmacophore modeling, to name a few examples.^{8–11} However, several methodologies, like conventional QSAR, make assumptions that are not necessarily valid and, thus, may present misleading results including non-predictive models.¹² For example, one common assumption is that a lead series of compounds has a common binding mode or mechanism of action.^{13,14} For this reason, understanding the activity landscape and early detection of activity cliffs¹⁵ can be crucial to the success of computational models.¹⁶

The SAR of a data set can be conceptualized as an *activity landscape* where biological activity adds another dimension to the chemical space.¹⁷ The activity landscape has been compared to rolling hills or continuous SAR where small changes in molecular structure are associated with small changes in activity.¹⁵ A discontinuous SAR or rugged activity landscape, however, is populated with molecules with small changes in structure but large changes in activity (*activity cliffs*).¹⁵ Such landscapes are common in lead optimization. It can also happen that structurally diverse compounds have similar activity, which is the basis of scaffold hopping.¹⁸ Additionally, active regions with wide variations in chemical structure but small variations in biological activity may

* Corresponding author. Tel.: +1 772 345 4685; fax: +1 772 345 3649.

E-mail address: jmedina@tpims.org (J.L. Medina-Franco).

suggest different binding modes or sites, or may reveal the effect of additional mechanisms such as the interaction with membranes that are not typically considered in several modeling approaches.¹⁶ Understanding the activity landscape of a data set is, however, a difficult task because the landscape may be highly complex involving a combination of smooth and rugged regions.¹⁴ Another major challenge is the dependence of chemical space on molecular representation.^{19,20}

Different approaches are emerging to characterize systematically the activity landscape and are reviewed in Bajorath et al.¹⁷ These include the structure–activity relationship index (SARI),²¹ structure-landscape index (SALI),^{16,22} structure–activity similarity (SAS) maps²³ and network-like similarity graphs (NSG).²⁴ SARI and NSG have been used to detect molecules with small changes in structure but large changes in selectivity (*selectivity cliffs*).²⁵ We proposed using multiple structural representations to derive a consensus model for the activity landscape and identify consensus activity cliffs.²⁶ Recently, 3D representations of the activity landscape were proposed, confirming the existence of consensus activity cliffs and representation-dependent cliffs.²⁷

Herein we conducted a comprehensive analysis of the activity landscape of 32 benzimidazoles mostly synthesized and tested in our group against *T. vaginalis* and *G. intestinalis* (Table 1). Compounds are non 2-methylcarbamates and their mechanism of action remains unknown.^{28,29} The analysis was based on pairwise comparisons of the activity similarity and molecular similarity. For each parasite, pairwise SAR was visually depicted using SAS maps. Quantitative analyses of the SAS maps were used to identify

consensus activity cliffs and develop consensus models of the activity landscape. We also compared the SAR of the benzimidazoles between the two parasites.

2. Methods

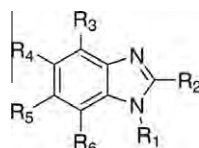
2.1. Data set

The chemical structure of 32 previously reported benzimidazoles^{4–7} is presented in Table 1 along with the biological activity against *T. vaginalis* and *G. intestinalis*. Table 1 lists the 50% inhibitory concentration (IC₅₀) in *in vitro* susceptibility assays for each parasite as pIC₅₀ (–log IC₅₀). For *T. vaginalis*, the activity ranges from 2 nM (pIC₅₀ = 8.7) to 29,512 nM (pIC₅₀ = 4.53); median pIC₅₀ = 6.7. For *G. intestinalis*, the activity ranges from 22.9 nM (pIC₅₀ = 7.64) to 10,471 nM (pIC₅₀ = 4.98); median pIC₅₀ = 7.0. The activity of the 32 compounds was obtained by the same group under similar conditions.

2.2. 2D and 3D structural similarity

For each pair of molecules m_i and m_j , pairwise *structural similarities* (SS_{ij}) were computed using the Tanimoto coefficient^{30,31} with the following 2D molecular representations as implemented in molecular operating environment (MOE):³² MACCS keys (166 bits), pharmacophore graph triangle (i.e., graph-based three point pharmacophores) (GpiDAPH3), typed graph distance (TGD), and typed graph triangle (TGT). We also used the following 2D (32-bit)

Table 1
Chemical structures of benzimidazoles and biological activity against *T. vaginalis* and *G. intestinalis*



	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	pIC ₅₀ <i>T. vaginalis</i>	pIC ₅₀ <i>G. intestinalis</i>	ΔpIC ₅₀	Ref.
1	H	CF ₃	H	H	H	H	5.50	6.97	1.47	4
2	CH ₃	CF ₃	H	CF ₃	H	H	5.39	5.94	0.55	4
3	CH ₃	CF ₃	H	H	CF ₃	H	5.27	5.05	-0.22	4
4	CH ₃	CF ₃	H	Propylthio	H	H	6.70	4.98	-1.72	5
5	CH ₃	CF ₃	H	H	Propylthio	H	5.59	5.85	0.26	5
6	CH ₃	CF ₃	H	Benzoyl	H	H	4.53	5.96	1.43	5
7	CH ₃	CF ₃	H	H	Benzoyl	H	4.97	5.89	0.92	5
8	H	CF ₃	H	Br	Br	H	6.66	6.92	0.26	6
9	H	CF ₃	Br	Br	Br	Br	8.70	7.25	-1.45	6
10	H	C ₂ F ₅	H	Cl	Cl	H	6.52	6.25	-0.27	6
11	H	CF ₃	H	NO ₂	NO ₂	H	6.24	6.62	0.38	6
12	H	C ₂ F ₅	Br	Br	Br	Br	5.00	7.64	2.64	6
13	CH ₃	CONH ₂	H	H	Cl	H	6.96	7.12	0.16	7
14	CH ₃	CONHCH ₃	H	H	Cl	H	6.98	7.15	0.17	7
15	CH ₃	CON(CH ₃) ₂	H	H	Cl	H	6.63	7.40	0.77	7
16	CH ₃	COOCH ₂ CH ₃	H	H	Cl	H	7.72	7.32	-0.4	7
17	CH ₃	CONH ₂	H	Cl	H	H	6.73	6.63	-0.1	7
18	CH ₃	CONHCH ₃	H	Cl	H	H	6.45	6.45	0	7
19	CH ₃	CON(CH ₃) ₂	H	Cl	H	H	6.68	6.61	-0.07	7
20	CH ₃	COOCH ₂ CH ₃	H	Cl	H	H	7.57	7.40	-0.17	7
21	CH ₃	CONH ₂	H	Cl	Cl	H	6.87	6.34	-0.53	7
22	CH ₃	CONHCH ₃	H	Cl	Cl	H	6.65	6.82	0.17	7
23	CH ₃	CON(CH ₃) ₂	H	Cl	Cl	H	7.12	7.13	0.01	7
24	CH ₃	COOCH ₂ CH ₃	H	Cl	Cl	H	7.53	7.56	0.03	7
25	CH ₃	CONH ₂	H	H	H	H	6.78	7.03	0.25	7
26	CH ₃	CONHCH ₃	H	H	H	H	6.98	7.22	0.24	7
27	CH ₃	CON(CH ₃) ₂	H	H	H	H	6.37	6.29	-0.08	7
28	CH ₃	COOCH ₂ CH ₃	H	H	H	H	7.07	7.16	0.09	7
29	CH ₃	COCH ₃	H	H	H	H	6.68	7.06	0.38	7
30	CH ₃	COCH ₃	H	Cl	H	H	6.88	7.30	0.42	7
31	CH ₃	COCH ₃	H	H	Cl	H	6.64	7.17	0.53	7
32	CH ₃	COCH ₃	H	Cl	Cl	H	7.20	7.46	0.26	7

fingerprints as implemented in Canvas:^{33,34} radial (also known as extended connectivity fingerprints), dendritic, atom pairs and MOLPRINT 2D. To compute 3D similarities, a single low-energy conformation was considered for each molecule obtained with geometry optimization using the PM3 semiempirical method as implemented in Spartan'02.^{35,36} 3D similarity values were calculated with the MOE pharmacophore atom triangle (piDAPH3) and pharmacophore atom quadruplet (piDAPH4) fingerprints, and Canvas 3- and 4-point pharmacophores. Despite the inherent conformational issues, the use of 3D structural representations were valuable to characterize the activity landscapes (see below).

2.3. Property similarity

The following properties were computed with Canvas: molecular weight (MW), number of rotatable bonds (RB), hydrogen bond acceptors (HBA), hydrogen bond donors (HBD), polar surface area (PSA), and the octanol/water partition coefficient ($A \log P$). Properties were first auto-scaled with mean centering using the equation:

$$p_{ki} = \frac{P_{ki} - \bar{P}_k}{\sigma_{P_k}} \quad (1)$$

where p_{ki} denotes the scaled version of the k th property for the i th molecule, P_{ki} denotes the unscaled value, and \bar{P}_k and σ_{P_k} denote, respectively, the mean and standard deviation of the k th property over all molecules in the study.

The Euclidean distance between a pair of molecules was then computed with the expression:³¹

$$d_{ij} = [\sum_{k=1}^K (p_{ki} - p_{kj})^2]^{1/2} \quad (2)$$

where d_{ij} denotes the Euclidean distance between the i th and j th molecules, and p_{ki} and p_{kj} denote the value of the scaled property k of the i th and j th molecules, respectively. In this study, $K = 6$.

Euclidean distances were scaled from 0 to 1 as follows:

$$sd_{ij} = \frac{d_{ij} - \min d_{ij}}{\max d_{ij} - \min d_{ij}} \quad (3)$$

where sd_{ij} is the scaled distance, and $\max d_{ij}$ and $\min d_{ij}$ indicate the range of distances in the data set. Pairwise property similarities were measured with the expression:

$$PS_{ij} = 1 - sd_{ij} \quad (4)$$

where PS_{ij} is the *property similarity* of the i th and j th molecules, and sd_{ij} is the scaled distance.

2.4. Activity similarity

For each pair of molecules the activity similarity for *T. vaginalis* and *G. intestinalis* was measured as follows:

$$AS_{i,j} = 1 - \frac{|A_i - A_j|}{\max - \min} \quad (5)$$

where A_i and A_j are the activities of the i th and j th molecules (pIC_{50} values) and \max - \min indicate the range of activities in the data set.

2.5. Activity landscape with SAS maps

For each target parasite, SAS maps²³ were generated by plotting the activity similarity against the structural similarity for each pair of compounds. A general form of the SAS map is presented in Figure 1. In this map the activity similarity is represented in the Y-axis and molecular similarity is plotted in the X-axis. A variant of the SAS maps represents potency difference in the Y-axis.^{26,27}

SAS maps provide a visual and quantitative characterization of the activity landscape.¹⁷ Four zones can be distinguished in Figure 1, labeled as regions I–IV. Data points that fall into region I correspond to pairs of molecules with high activity similarity and low molecular similarity and therefore are associated with regions of scaffold hopping. If the compounds in the data set share the same core scaffold and the differences are only in the attachment points, then region I is associated with *side chain hopping*.³⁷ Points plotted in region II denote pairs of molecules with high molecular similarity and high activity similarity. Thus, compounds in this region are in a smooth or continuous SAR landscape. Data points in region III denote pairs of molecules with low molecular similarity and low activity similarity. Region IV identifies pairs of molecules that have high molecular similarity and low activity similarity and therefore correspond to activity cliffs or discontinuous SAR. Data points that are consistently put in the same region by a number of molecular representations contribute to defining a *consensus model* of the activity landscape.

In order to characterize the SAS maps obtained with different similarity measures, each map was partitioned by imposing activity and molecular similarity thresholds along the Y- and X-axis, respectively, and then counting the number of data pairs in the resultant regions I–IV. A similar strategy was recently employed to successfully characterize potency difference vs. structure similarity plots.²⁶ In this study, two activity similarity thresholds were investigated, namely, 0.5 and 0.75, corresponding approximately to 1 and 2 log units in potency difference for *T. vaginalis*, and 0.7 and 1.4 log units for *G. intestinalis*. For similarity, the median similarity of the most active compounds in the data set was considered. For *T. vaginalis*, seven compounds with $pIC_{50} \geq 7.00$ ($IC_{50} \leq 100$ nM) were regarded as active (9, 16, 20, 23, 24, 28 and 32). For *G. intestinalis*, also seven compounds with $pIC_{50} \geq 7.30$ ($IC_{50} \leq 50$ nM) were regarded as active (12, 15, 16, 20, 24, 30 and 32). The reason to use slightly different thresholds of pIC_{50} was twofold; in order to select the same number of actives for each parasite and because the median of the pIC_{50} values for *G. intestinalis* is greater than the median for *T. vaginalis* (see above). Note, however, that other thresholds for activity could be applied. Since different molecular representations lead to different ranges of similarity values for the same set of compounds, the threshold for the structural similarity depends on the representation used. Therefore, normalizing to the median similarity of a descriptor allowed us the ability to compare between different descriptor sets.²⁶

To further compare the SAS maps obtained from different structural similarity methods, we used the *Degree of Consensus* (DoC) introduced in an earlier study.²⁶ DoC measures the number of data

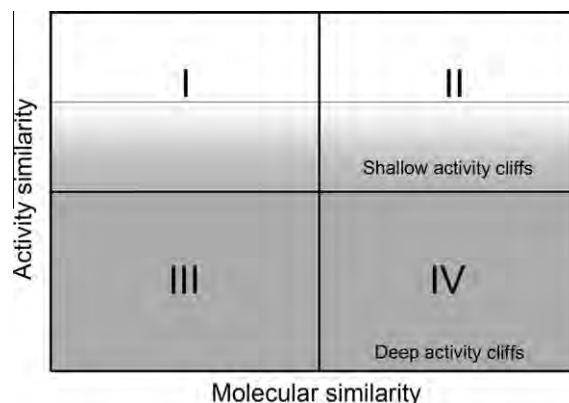


Figure 1. General form of the structure–activity similarity (SAS) map showing four major regions. Regions I and II contain data pairs for scaffold hopping and smooth SAR, respectively. Region IV indicates discontinuous SAR and activity cliffs. Regions of deep and shallow activity cliffs are also shown. See text for details.

points consistently put into the same region and was computed with the following expression:

$$\text{DoC}_{m,n}^R = \frac{Cp_{m,n}}{p_m + p_n - Cp_{m,n}} \quad (6)$$

where $Cp_{m,n}$ is the number of Consensus Pairs in region R ($R = \text{I–IV}$) between methods m and n ; p_m is the number of pairs of molecules assigned by method m in region R , and p_n is the number of pairs of molecules assigned by method n in the same region. To note, DoC depends on the thresholds used to define each region. Results were summarized as DoC matrices.

2.6. Consensus SAS maps

To develop consensus models of the activity landscapes, we combined structural similarities obtained with uncorrelated representations into a single similarity measure. There are a number of ways to combine similarity values.³⁸ In this work we computed the mean similarity of four orthogonal fingerprints for this data set (radial, MACCS keys, TGD and piDAPH3), but other measures can be explored. Property similarity was not considered in order to obtain the mean. Consensus SAS maps for each parasite were generated by plotting the activity similarity against the mean fingerprint similarity.

3. Results and discussion

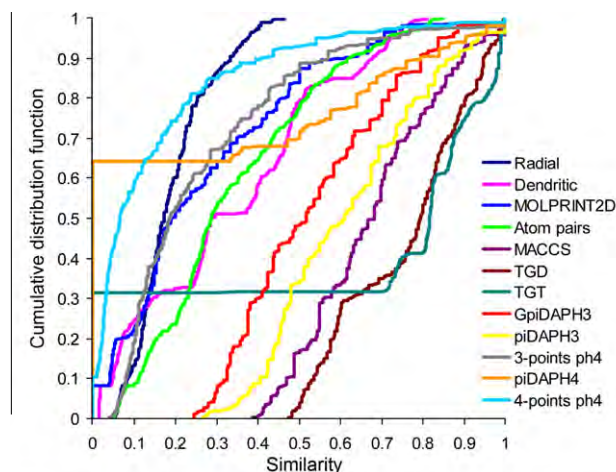
3.1. Distribution of fingerprint similarity measures

The 496 pairwise similarities of the 32 benzimidazoles calculated with the 12 fingerprint-based structural representations are summarized in Figure 2 as cumulative distribution functions (CDF). The table at the bottom of the figure summarizes the statistics of each curve indicating the maximum, third and first quartile, median, mean, and standard deviation.

2D and 3D fingerprints showed a wide variation of distributions. 2D fingerprints with the highest similarity values were TGD, MACCS keys and GpiDAPH3 which had median values of 0.80, 0.68 and 0.52, respectively, and comparable standard deviation (0.16–0.18). 2D fingerprints with the lowest similarity values were dendritic, atom pairs, MOLPRINT 2D and radial. Concerning the 3D fingerprints, the spatial three-point pharmacophore piDAPH3 had a nearly normal distribution. In contrast, similarity values obtained with piDAPH4, 3- and 4-point pharmacophores showed non-normal distributions (as can be deduced from the non-sigmoidal shape of the corresponding CDF). Despite the fact that the 32 molecules share the benzimidazole scaffold, it was noteworthy that several fingerprints were able to differentiate the compounds, thus identifying activity cliffs (see below).

3.2. Correlation between molecular similarities

The correlation between 2D and 3D fingerprint representations for the 496 pairwise similarities is shown in Table 2. The correlation matrix shows the relationships between the 12 fingerprints. Additionally, the matrix shows the relationship between the fingerprint, property similarity, and activity similarity. Very high correlations between 2D methods occur for radial and dendritic; MOLPRINT 2D and atom pairs; radial and atom pairs; dendritic and atom pairs; TGD and TGT (correlation ≥ 0.92). Other high correlations between 2D methods are atom pairs and MACCS (0.89); MOLPRINT 2D and MACCS (0.84); atom pairs and GpiDAPH3 (0.83). High correlations between 3D fingerprints occur for Canvas 3- and 4-point pharmacophores (0.97), and between piDAPH3 and 3-point pharmacophore (0.82). Comparing the correlation between 2D and 3D fingerprints,



	Max	Q3	Median	Q1	Min	Mean	STD
Radial	0.46	0.24	0.17	0.13	0.05	0.19	0.09
Dendritic	0.81	0.48	0.29	0.11	0.02	0.33	0.23
MOLPRINT 2D	1.00	0.42	0.20	0.12	0.00	0.27	0.22
Atom pairs	0.85	0.49	0.29	0.21	0.05	0.34	0.20
MACCS	1.00	0.80	0.68	0.55	0.39	0.68	0.16
TGD	1.00	0.88	0.80	0.59	0.47	0.76	0.16
TGT	1.00	0.90	0.82	0.00	0.00	0.60	0.41
GpiDAPH3	1.00	0.69	0.52	0.38	0.25	0.54	0.18
piDAPH3	1.00	0.76	0.61	0.46	0.27	0.62	0.18
3-points ph4	1.00	0.37	0.19	0.11	0.04	0.27	0.21
piDAPH4	1.00	0.58	0.00	0.00	0.00	0.24	0.34
4-points ph4	1.00	0.20	0.07	0.03	0.00	0.15	0.20

Figure 2. Cumulative distribution functions of 496 pairwise structural similarities using different 2D and 3D fingerprint representations. The table summarizes the information of the distributions. Q3 and Q1 indicate the third and first quartile, respectively.

the highest correlation was for GpiDAPH3 and piDAPH3 (0.94). The correlation between property similarity and fingerprint similarity ranges between 0.44 (piDAPH4) and 0.78 (atom pairs and MACCS). The matrix also shows a low correlation between any of the molecular representations with activity similarity for *T. vaginalis* (correlation ≤ 0.48) or *G. intestinalis* (≤ 0.30). The correlation between activity similarities for *T. vaginalis* and *G. intestinalis* was low (0.31) indicating the presence of pairs of compounds with different effects against the two parasites (see below).

Despite the high correlations between several 2D and 3D fingerprints, we selected as many orthogonal fingerprint representations as possible to characterize the landscapes. Thus, we selected radial, MACCS, TGD (2D), and piDAPH3 (3D). The maximum correlation between any of these five selected fingerprints was 0.78 (radial and MACCS), and the minimum correlation was 0.69 (TGD and piDAPH3). In addition, we employed property similarities as described below.

3.3. Activity landscape

3.3.1. SAS maps

Figures 3 and 4 depict the SAS maps for *T. vaginalis* and *G. intestinalis*, respectively. The maps show the relationship between activity similarity and molecular similarity obtained with four selected fingerprints and property similarity. Each plot contains 496 data points that represent a pairwise comparison. Data points were further distinguished by the activity of the most active compound in the pair in a continuous scale from green (least active) to red (most active). It is also possible to generate a visual representation of the plots coloring only data points where at least one compound in the pair is active (Figs. S1 and S2 in Supplementary data).²⁶

Table 2
Correlation matrix for the pairwise activity, property and structure similarities

	Radial	Dendritic	MOLPRINT 2D	Atom pairs	MACCS	TGD	TGT	GpiDAPH3	piDAPH3	3-point ph4	piDAPH4 4-point ph4	Properties AS	AS <i>T. vaginalis</i>	AS <i>G. intestinalis</i>	
Radial	1.00														
Dendritic	0.94	1.00													
MOLPRINT 2D	0.88	0.86	1.00												
Atom pairs	0.92	0.92	0.93	1.00											
MACCS	0.78	0.81	0.84	0.89	1.00										
TGD	0.71	0.82	0.69	0.76	0.77	1.00									
TGT	0.69	0.82	0.63	0.71	0.73	0.93	1.00								
GpiDAPH3	0.79	0.76	0.77	0.83	0.76	0.69	0.56	1.00							
piDAPH3	0.76	0.73	0.74	0.79	0.73	0.69	0.57	0.94	1.00						
3-point ph4 ^a	0.78	0.75	0.81	0.83	0.77	0.68	0.55	0.82	0.81	1.00					
piDAPH4	0.40	0.36	0.43	0.48	0.53	0.32	0.19	0.61	0.62	0.49	1.00				
4-point ph4 ^a	0.73	0.69	0.78	0.79	0.72	0.63	0.49	0.78	0.77	0.97	0.50	1.00			
Properties	0.69	0.69	0.74	0.78	0.78	0.65	0.59	0.73	0.74	0.75	0.44	0.69	1.00		
AS ^b <i>T. vaginalis</i>	0.43	0.40	0.44	0.45	0.28	0.26	0.19	0.48	0.47	0.42	0.25	0.39	0.37	1.00	
AS ^b <i>G. intestinalis</i>	0.28	0.20	0.28	0.25	0.15	-0.01	-0.05	0.29	0.30	0.27	0.17	0.26	0.20	0.31	1.00

^a ph4: pharmacophore

^b AS: activity similarity

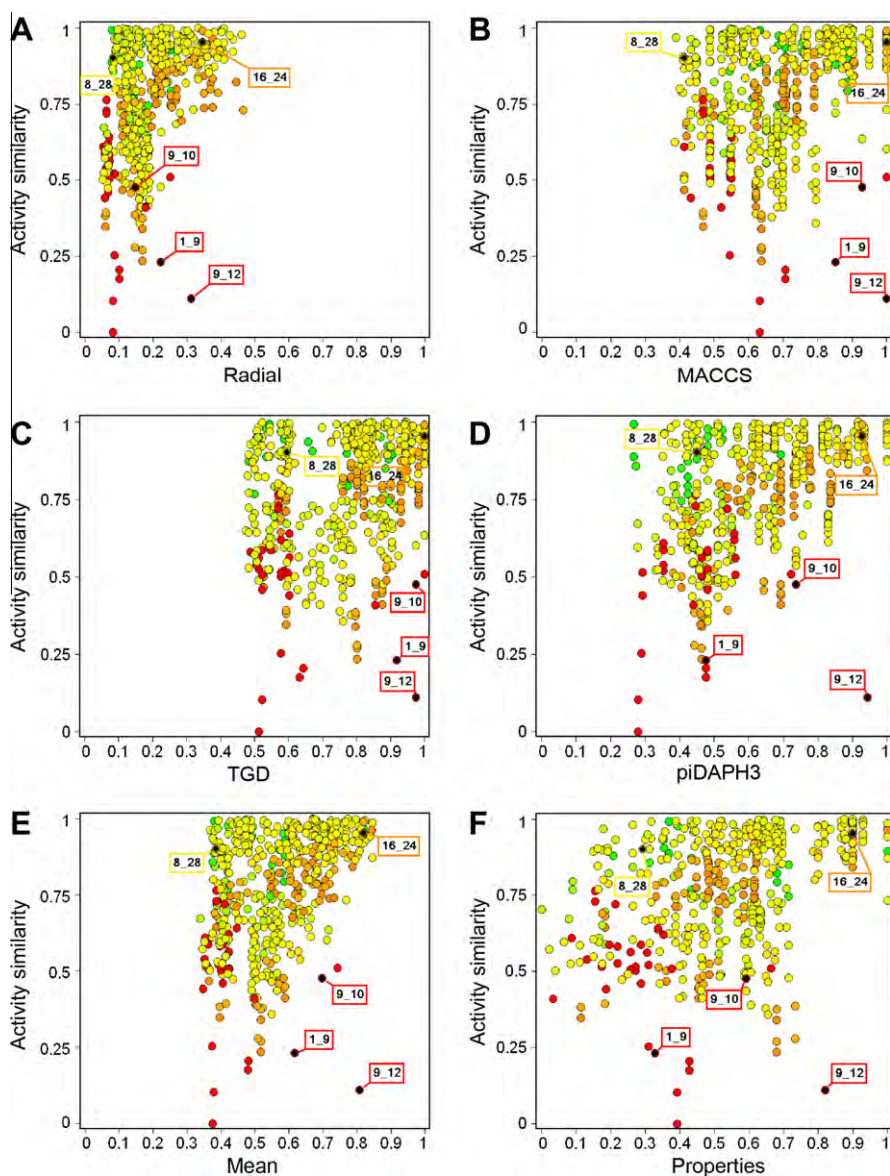


Figure 3. SAS maps for *T. vaginalis* with different structural representations. Each data point indicates a pairwise comparison of 32 benzimidazoles (496 data points). Data points are color-coded by the activity of the most active compound in the pair using a continuous scale from green (less active) to red (more active). Each panel corresponds to a different structural representation: (A) Radial; (B) MACCS keys; (C) TGD; (D) piDAPH3; (E) mean fingerprint similarity, and (F) properties. Selected pairs are marked in black and labeled with the compound numbers.

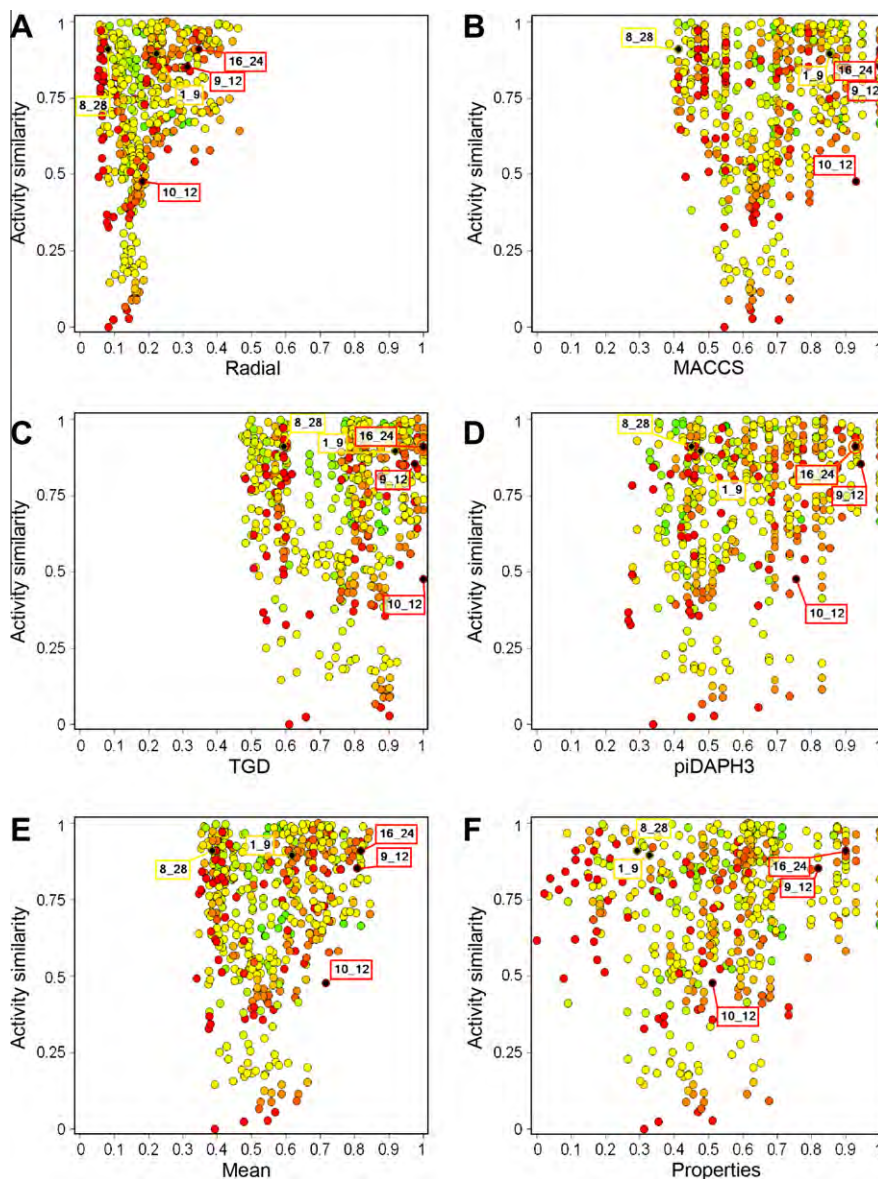


Figure 4. SAS maps for *G. intestinalis* with different structural representations. Each data point indicates a pairwise comparison of 32 benzimidazoles (496 data points). Data points are color-coded by the activity of the most active compound in the pair using a continuous scale from green (less active) to red (more active). Each panel corresponds to a different structural representation: (A) Radial; (B) MACCS keys; (C) TGD; (D) piDAPH3; (E) mean fingerprint similarity, and (F) properties. Selected pairs are marked in black and labeled with the compound numbers.

As expected from the CDF in Figure 2, similarity values obtained with radial fingerprints are shifted to the low similarity value range along the X-axis (<0.46) while the similarity values for MACCS, TGD and piDAPH3 are more spread out. Interestingly, there are several pairs in Figures 3 and 4 with similarities of 1.0 (21 pairs for MACCS keys, 15 for TGD, 18 for piDAPH3 and 8 for property similarity). A number of these points correspond to positional isomers or compounds with different halogen substitution pattern (discussed below). Notably, radial fingerprints can distinguish all 496 pairs due to its high resolution. This is an expected behavior since radial fingerprints were designed for structure activity studies in contrast with topological fingerprints that were developed for substructure and similarity searching.^{34,39}

As described in the Section 2, the four regions of the SAS maps (I–IV in Fig. 1) can also be identified in Figures 3 and 4. Data pairs in regions I and II are located in a continuous SAR while pairs of molecules in region IV represent activity cliffs. As discussed above, the boundary between regions I/II and III/IV depends on the molec-

ular representation used. However, it is possible to detect pairs of compounds that are located in the same *relative* region of each map, that is, *consensus pairs*. Figures 3 and 4 show examples of consensus pairs in regions I (compounds 8 and 28) and II (16 and 24) for *T. vaginalis* and *G. intestinalis*. Figure 3 also shows a consensus pair in region IV (9 and 12) for *T. vaginalis*.

3.3.2. Quantitative characterization of the SAS maps

In order to conduct a systematic and quantitative analysis of the data obtained in this study, the SAS maps were divided into four quadrants (regions I–IV in Fig. 1) by defining thresholds for activity similarity and molecular similarity (see Section 2). Table 3 indicates the median similarity of the active compounds and the number of data pairs that can be found in regions I–IV for different molecular representations. Table 3 also summarizes the number of pairs with at least one active molecule or *active pair*. Interestingly, the median fingerprint similarity of the actives for *T. vaginalis* and *G. intestinalis* (Table 3) is slightly higher than the median similarity of the 32 compounds (Fig. 2). Data in Table 3 can be visualized

Table 3
Distribution of data points in different regions of the SAS maps

Representation (AS threshold) ^a	Median similarity of actives ^b	I		II		III		IV	
		Total ^c	Active pairs ^d	Total	Active pairs	Total	Active pairs	Total	Active pairs
<i>T. vaginalis</i> (0.5)									
Radial	0.22	286	94	160	66	47	34	3	2
MACCS	0.71	271	102	175	58	42	31	8	5
TGD	0.95	401	138	45	22	48	34	2	2
piDAPH3	0.78	346	129	100	31	49	35	1	1
Properties	0.62	285	120	161	40	38	25	12	11
<i>T. vaginalis</i> (0.75)									
Radial	0.22	145	50	149	59	188	78	14	9
MACCS	0.71	142	59	152	50	171	74	31	13
TGD	0.95	254	89	40	20	195	83	7	4
piDAPH3	0.78	204	80	90	29	191	84	11	3
Properties	0.62	160	74	134	35	163	71	39	16
<i>G. intestinalis</i> (0.5)									
Radial	0.24	301	83	114	50	81	38	0	0
MACCS	0.79	286	91	129	42	76	35	5	3
TGD	0.93	327	90	88	43	80	37	1	1
piDAPH3	0.84	356	113	59	20	81	38	0	0
Properties	0.61	241	87	174	46	63	26	18	12
<i>G. intestinalis</i> (0.75)									
Radial	0.24	184	54	76	35	198	67	38	15
MACCS	0.79	171	58	89	31	191	68	45	14
TGD	0.93	199	57	61	32	208	70	28	12
piDAPH3	0.84	216	72	44	17	221	79	15	3
Properties	0.61	138	54	122	35	166	59	70	23

^a Regions I–IV are defined by the median similarity of active compounds and a threshold of activity similarity. Two thresholds of activity similarity (AS) were investigated 0.5 and 0.75 (see also Fig. 1 and text for details).

^b Median similarity of compounds **9, 16, 20, 23, 24, 28, 32** for *T. vaginalis*, and **12, 15, 16, 20, 24, 30, 32** for *G. intestinalis*.

^c Total number of data points (pair of compounds) in the region.

^d Number of data points with at least one active compound in the pair.

using pie charts (Fig. S3 in Supplementary data illustrates examples of binned SAS maps²⁶).

Considering an activity similarity threshold of 0.5 for *T. vaginalis* and *G. intestinalis*, most pairs of compounds were in region I, and the frequency decreased in the order I > II > III > IV (Table 3). This occurred for most fingerprint representations and property similarity.⁴⁰ A similar result was obtained for the number of active pairs.

Not surprisingly, considering an activity similarity threshold of 0.75, the total number of pairs and active pairs in regions I and II decreased, and the number of pairs in regions III and IV increased (as compared to the populations at a threshold of 0.5). This was observed for all representations and for both parasites (Table 3). However, for *G. intestinalis*, most of the pairs were in region III for all molecular representations (and most representations for *T. vaginalis*). Similar results were obtained for the number of active pairs. In general, at the activity similarity threshold of 0.75, the frequency of total pairs and active pairs decreased in the order III > I > II > IV (Table 3).

3.3.3. Deep and shallow activity cliffs

The number of activity cliffs in the data set depends on the criteria used to define a pair of compounds as similar. It follows that activity cliffs can be classified further using, for example, different thresholds of activity similarity. In this work we define a *deep activity cliff* if the pair of similar compounds have 'a large' difference in activity (activity similarity ≤ 0.5), and define a *shallow activity cliff* if the difference in activity is smaller ($0.5 < \text{activity similarity} \leq 0.75$).²⁷ This is schematically illustrated in Figure 1. Different activity similarity thresholds can be used to define deep and shallow cliffs. Moreover, the degree of molecular similarity can be used also to define deep or shallow activity cliffs. According to these definitions, there were between one and eight

deep activity cliffs in the landscape of *T. vaginalis*, depending on the fingerprint representation. The number of deep activity cliffs for *G. intestinalis* was lower, between zero and five (i.e., the number of deep activity cliffs equals the number of pairs in region IV at activity threshold of 0.5, Table 3). Considering molecular properties as molecular representation, the number of deep cliffs for *T. vaginalis* and *G. intestinalis* was 12 and 18, respectively. Noteworthy, the numbers of deep and shallow cliffs, as defined in this work, are relative to the data set. This is because deep and shallow cliffs are defined based on activity similarity that depends on the activity range of the data set (Eq. 5).

The number of shallow activity cliffs can be calculated from Table 3 taking the difference between the number of total pairs in region IV at activity similarity thresholds of 0.75 and 0.5. For example, for *T. vaginalis*, there are $14 - 3 = 11$ shallow cliffs considering radial fingerprints and $31 - 8 = 23$ shallow cliffs considering MACCS keys. Table S1 in the Supplementary data summarizes the number of shallow activity cliffs for each parasite and each molecular representation. Noteworthy, for all molecular representations, the number of shallow cliffs for *G. intestinalis* was higher than the number of shallow cliffs for *T. vaginalis*. Examples of deep and shallow activity cliffs are discussed below.

Results above, and in previous studies, support the importance of considering several representations^{26,27} and lead to the following questions: are there *consensus pairs*?²⁶ Is it possible to derive a consensus model of the activity landscape for a given data set?

3.3.4. Degree of consensus

Despite the mid-to-low correlations between the molecular representations (Table 2) and the different distributions of pairwise similarity values (Fig. 2), it was possible to find a number of consensus pairs in different regions of the landscape. The number of pairs of compounds that two methods put into the same region,

T. vaginalis (Activity similarity threshold 0.5)

Region I					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.81	1.00			
TGD	0.67	0.62	1.00		
piDAPH3	0.77	0.73	0.80	1.00	
Properties	0.72	0.79	0.67	0.74	1.00

Region III					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.85	1.00			
TGD	0.94	0.88	1.00		
piDAPH3	0.96	0.86	0.98	1.00	
Properties	0.77	0.70	0.76	0.78	1.00

G. intestinalis (Activity similarity threshold 0.5)

Region I					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.77	1.00			
TGD	0.75	0.76	1.00		
piDAPH3	0.77	0.80	0.87	1.00	
Properties	0.67	0.77	0.68	0.68	1.00

Region III					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.94	1.00			
TGD	0.99	0.95	1.00		
piDAPH3	1.00	0.94	0.99	1.00	
Properties	0.78	0.76	0.77	0.78	1.00

T. vaginalis (Activity similarity threshold 0.75)

Region I					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.74	1.00			
TGD	0.52	0.49	1.00		
piDAPH3	0.69	0.66	0.74	1.00	
Properties	0.65	0.79	0.58	0.70	1.00

Region III					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.89	1.00			
TGD	0.93	0.87	1.00		
piDAPH3	0.90	0.85	0.93	1.00	
Properties	0.80	0.78	0.80	0.80	1.00

G. intestinalis (Activity similarity threshold 0.75)

Region I					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.76	1.00			
TGD	0.76	0.75	1.00		
piDAPH3	0.76	0.78	0.86	1.00	
Properties	0.62	0.75	0.65	0.64	1.00

Region III					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.84	1.00			
TGD	0.84	0.83	1.00		
piDAPH3	0.86	0.86	0.92	1.00	
Properties	0.76	0.79	0.74	0.75	1.00

Region II					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.71	1.00			
TGD	0.20	0.17	1.00		
piDAPH3	0.52	0.49	0.28	1.00	
Properties	0.55	0.68	0.20	0.47	1.00

Region IV					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.22	1.00			
TGD	0.25	0.25	1.00		
piDAPH3	0.33	0.13	0.50	1.00	
Properties	0.15	0.18	0.08	0.08	1.00

Region II					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.52	1.00			
TGD	0.39	0.44	1.00		
piDAPH3	0.34	0.45	0.50	1.00	
Properties	0.45	0.63	0.42	0.34	1.00

Region IV					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	NA				
MACCS	0.00	1.00			
TGD	0.00	0.20	1.00		
piDAPH3	NA	0.00	0.00	NA	
Properties	0.00	0.10	0.00	0.00	1.00

Region II					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.75	1.00			
TGD	0.20	0.17	1.00		
piDAPH3	0.57	0.54	0.31	1.00	
Properties	0.63	0.78	0.23	0.56	1.00

Region IV					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.36	1.00			
TGD	0.24	0.19	1.00		
piDAPH3	0.14	0.17	0.13	1.00	
Properties	0.15	0.25	0.07	0.11	1.00

Region II					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.54	1.00			
TGD	0.44	0.49	1.00		
piDAPH3	0.38	0.48	0.54	1.00	
Properties	0.45	0.65	0.44	0.36	1.00

Region IV					
	Radial	MACCS	TGD	piDAPH3	Properties
Radial	1.00				
MACCS	0.43	1.00			
TGD	0.29	0.33	1.00		
piDAPH3	0.26	0.33	0.39	1.00	
Properties	0.37	0.46	0.27	0.21	1.00

Figure 5. Degree of consensus (DoC) matrices for each region. Each entry in the corresponding matrix represents the agreement between two methods to place a pair of compounds into the same region. DoC is computed with Eq. 6 using data in Figure S4.

that is, number of consensus pairs, were recorded for *T. vaginalis* and *G. intestinalis* at the two activity similarity thresholds. Results are summarized in Figure S4 in Supplementary data. For both parasites, we identified several consensus pairs in all regions of the landscape at the two thresholds. The only exception was region IV for *G. intestinalis* at a threshold of 0.5; we found only one consensus pair between MACCS and TGD and two consensus pairs between MACCS and property similarity (Fig. S4). Examples of consensus pairs are discussed later in this section. The DoC between

two methods is presented in Figure 5. DoC measures the number of consensus pairs between two methods scaled by the total number of pairs that the two methods put into the same region (see Section 2). For both parasites, at the two activity similarity thresholds, DoC has high values in region III (0.70–1.00) followed by region I (0.49–0.98). In contrast, DoC has low values in region IV, in particular for *G. intestinalis* at activity similarity threshold of 0.5 (0–0.2). This means that there was better agreement between the methods used to assign molecules to region III than in any other

Table 4
Examples of consensus pairs of compounds in the SAS maps

Pair	Activity similarity		Fingerprint similarity					Property similarity
	<i>T. vaginalis</i>	<i>G. intestinalis</i>	Radial	MACCS	TGD	piDAPH3	Mean (STD)	
4_28 ^a	0.911	0.180	0.138	0.612	0.848	0.561	0.540 (0.295)	0.443
8_28 ^a	0.902	0.910	0.082	0.413	0.594	0.450	0.385 (0.216)	0.291
10_28 ^a	0.868	0.658	0.080	0.413	0.618	0.487	0.400 (0.229)	0.348
21_28 ^a	0.952	0.692	0.188	0.617	0.805	0.620	0.557 (0.262)	0.503
28_32 ^{a,b}	0.969	0.887	0.190	0.698	0.945	0.693	0.631 (0.317)	0.603
7_12 ^b	0.993	0.342	0.079	0.632	0.544	0.266	0.380 (0.254)	0.369
16_20 ^{a,b}	0.964	0.970	0.375	1.000	1.000	1.000	0.844 (0.313)	1.000
16_24 ^{a,b}	0.954	0.910	0.346	1.000	1.000	0.928	0.818 (0.317)	0.900
16_28 ^{a,b}	0.844	0.940	0.351	0.905	0.949	0.941	0.786 (0.291)	0.900
15_16 ^{a,b}	0.739	0.970	0.351	0.708	0.983	0.836	0.719 (0.270)	0.616
16_19 ^{a,b}	0.751	0.733	0.253	0.708	0.983	0.836	0.695 (0.315)	0.616
16_30 ^{a,b}	0.799	0.992	0.269	0.791	0.931	0.836	0.707 (0.297)	0.615
16_31 ^{a,b}	0.741	0.944	0.375	0.791	0.931	0.836	0.733 (0.246)	0.615
8_9 ^a	0.511	0.876	0.250	1.000	1.000	0.723	0.743 (0.354)	0.664
1_9 ^a	0.233	0.895	0.222	0.852	0.918	0.476	0.617 (0.327)	0.327
9_10 ^a	0.477	0.624	0.148	0.929	0.975	0.737	0.697 (0.380)	0.591
9_12 ^a	0.113	0.853	0.313	1.000	0.975	0.944	0.808 (0.331)	0.819
10_12 ^b	0.635	0.477	0.182	0.929	1.000	0.756	0.717 (0.371)	0.511

^a Active pair for *T. vaginalis*.

^b Active pair for *G. intestinalis*.

region. In contrast, it was more difficult to identify consensus activity cliffs than to identify consensus pairs in continuous regions of the SAR. Note that DoC is dependent on the criteria used to define the four regions.

Table 4 lists several examples of consensus pairs in the three most informative regions (I–II and IV) of the SAS map of *T. vaginalis* and *G. intestinalis*. Table 4 also lists the molecular similarity for selected fingerprint representations, property similarity and activity similarity.

For *T. vaginalis*, several examples of consensus pairs with low structural similarity and high activity similarity (region I) involve the active compound **28** ($IC_{50} = 86$ nM) such as **4_28**, **8_28**, **10_28**, **21_28** and **32_28** (Table 4). Figure 6A shows a comparison of the chemical structures for selected pairs in region I along with the activity and molecular similarity measures. In this figure, concentric ovals indicate different degrees of structural similarity to **28**. For example, compounds **8** and **10** are less similar to **28** as compared to **4**, **21** and **32**. Interestingly, **8** and **10** are also less active than **4**, **21** and **32**. Since all compounds in the set have the same scaffold, pairs in region I can be considered examples of side chain hopping (see above). Noteworthy, several whole-molecule fingerprints used in this study were able to detect low similarity due to the side chain substitutions. For example, the similarity for the above mentioned pairs with the known ‘low resolution’ 166-bit MACCS keys³⁹ ranges between 0.698 (**32_28**) and 0.413 (**8_28** and **10_28**). In contrast, for the same pairs, the radial similarity ranges between 0.190 (**32_28**) and 0.080 (**10_28**).

Some of the pairs in region I of the landscape of *T. vaginalis* were also in region I of the landscape of *G. intestinalis*. Examples were **8_28** and **28_32** with activity similarity values of 0.910 and 0.887, respectively (Table 4 and Fig. 6A). These results suggest a similar SAR for both parasites. Notable exceptions were **4_28** and **7_12** which have low activity similarity for *G. intestinalis* (0.18 and 0.342, respectively) indicating that in some instances the same change in the structure of the benzimidazoles produces different effects in the activity of *T. vaginalis* and *G. intestinalis*.

We also identified several consensus pairs in region II of the landscape of *T. vaginalis*. To note, a number of these pairs, with high structure similarity and high activity similarity (>0.84), included the active compound **16** ($IC_{50} = 19$ nM). Examples are pairs **16_20**, **16_24** and **16_28** (Fig. 6B). Interestingly, several fingerprint representations including MACCS keys, TGD and piDAPH3 were

unable to distinguish the positional isomer **16_20** (similarity of 1.0). However, the radial fingerprint did differentiate this pair demonstrating the high resolution of this type of fingerprint.³⁹ All compounds in these pairs have an ethyl ester at R_2 and are in a smooth region of landscape; changes in the substitution pattern with chlorine at positions 5 and 6 of the benzimidazole scaffold produce small changes in the activity (activity similarity between 0.844 and 0.964).

The pairs **16_15**, **16_19**, **16_30** and **16_31** are also located in region II of the landscape of *T. vaginalis* (Fig. 6B). The structural similarity of **15**, **19**, **30** and **31** with respect to **16** decreases (as captured by several molecular representations), and the activity similarity also decreases (down to 0.739–0.799). These results, schematically illustrated in Figure 6B, are in agreement with the ‘similarity principle’⁴¹ and further illustrate the smooth SAR associated with region II. To note, none of these compounds have an ethyl ester at R_2 emphasizing the importance of this substituent in the activity against *T. vaginalis*.

Several pairs with high structural similarity and high activity similarity for *T. vaginalis* were also located in region II of the landscape of *G. intestinalis* as illustrated by the pairs containing compound **16** ($IC_{50} = 48$ nM, Table 4). These results indicate that, in general, substitution with ethyl ester at R_2 increases the activity against both parasites. Similarly, substitution with one or two chlorine atoms at positions 5 and 6 of the benzimidazole scaffold does not affect significantly the activity against either parasite.

3.3.5. Consensus deep and shallow activity cliffs, and apparent cliffs

The importance of activity cliffs has been discussed in the literature.^{15,17,19} Activity cliffs are valuable for detecting specific structural changes important for activity. Furthermore, consensus activity cliffs have been conceptualized as those cliffs that occur across different molecular representations.²⁶ Figure 6C depicts examples of consensus cliffs (region IV) for different molecular representations.

For *T. vaginalis*, we found only one consensus deep cliff, pair **9_12** (cliff for all molecular representations with at least two log units in potency difference, Table 4). The structural difference between **9** and **12** is a CF_2 group at R_2 (CF_3 vs CF_2-CF_3). This change in structure produces a dramatic decrease in activity against *T. vaginalis* from $IC_{50} = 2$ nM (**9**) to 10,100 nM (**12**). It is anticipated that either

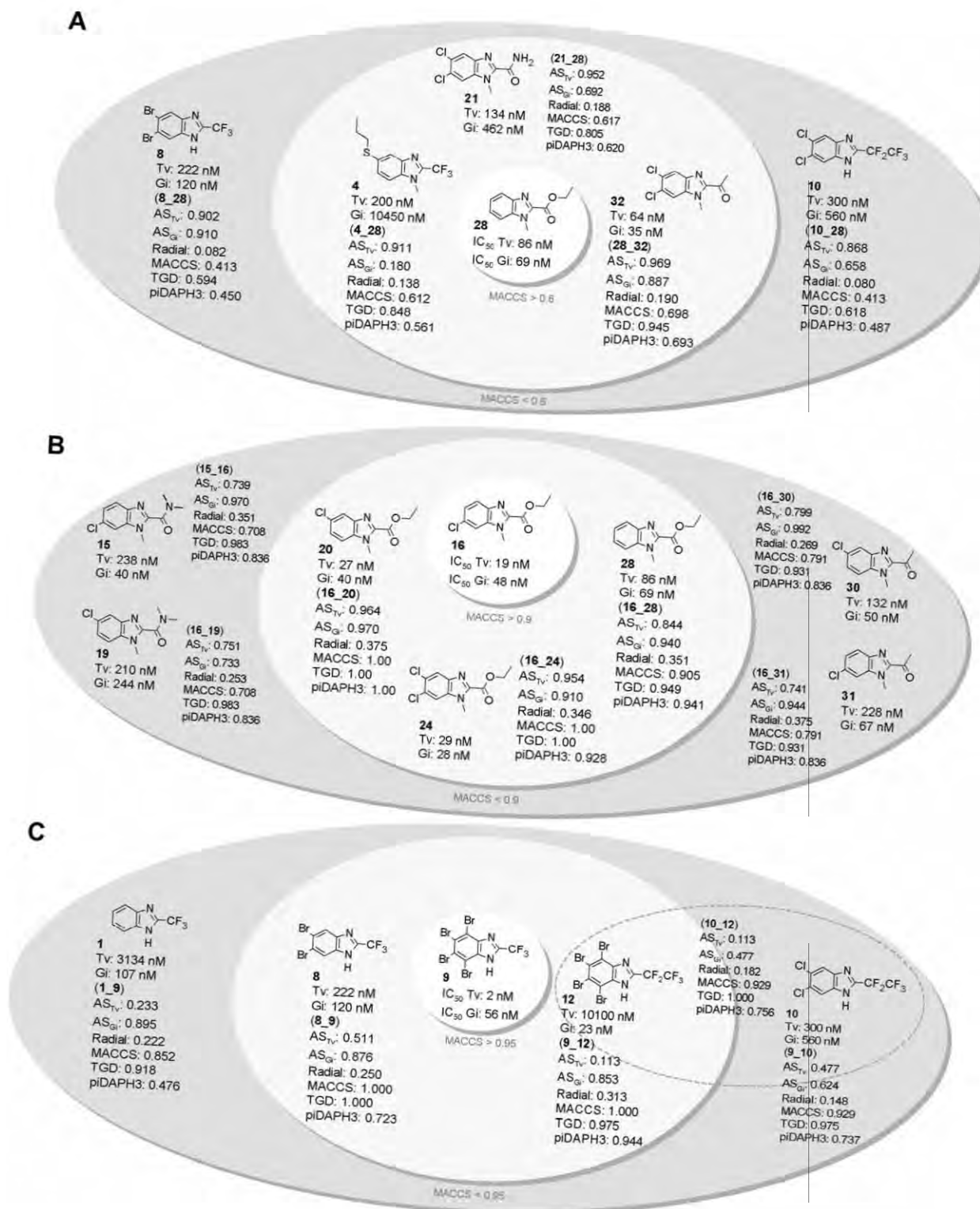


Figure 6. Representative consensus pairs in the activity landscapes of *T. vaginalis* and *G. intestinalis* arranged in concentric similarity ovals. Compounds in the inner ring are more structurally similar to the reference (center) than compounds in the outer ring: (A) region I, side chain hopping; (B) region II, smooth SAR, and (C) activity cliffs.

of these two compounds or both would be apparent outliers in a QSAR study.¹⁵ Interestingly, **9_12** was not an activity cliff for *G. intestinalis* (IC₅₀ = 56 nM vs 23 nM, respectively). A second consensus deep cliff in the landscape of *T. vaginalis* was the pair **1_9**, detected by radial and MACCS keys only. This is an example of an *apparent cliff* conceptualized as cliff identified just for some molecular representations.²⁶ One more example of an apparent cliff is

the pair **9_10** identified by MACCS keys and TGD. A borderline case between deep and shallow cliff for *T. vaginalis* is **8_9** (activity similarity value of 0.511) identified as cliff by radial, MACCS and TGD (but not by piDAPH3). Note, however, that MACCS and TGD could not distinguish this pair (similarity = 1).

Few consensus activity cliffs were identified in the landscape of *G. intestinalis*. For example, the pair **10_12** (Table 4 and Figs. 4 and

6C) was identified as a deep cliff by MACCS and TGD only (e.g., apparent cliff). As discussed above, when the activity similarity threshold is set to less restrictive changes in potency difference, the number of activity cliffs increases. An example of a shallow (and also apparent) cliff was the pair **9_10** identified by MACCS and TGD. This pair was also identified as an apparent cliff in the landscape of *T. vaginalis* (Table 4).

The presence of pairs of compounds in continuous and discontinuous regions of the landscape for *T. vaginalis* and *G. intestinalis* revealed the heterogeneous SAR for both parasites. Heterogeneous SARs has been reported for other activity classes.^{14,26,27} The landscape of *T. vaginalis* is characterized by the presence of two consensus deep activity cliffs and several data points in continuous regions of the SAR. In contrast, the landscape of *G. intestinalis* did not show consensus deep activity cliffs but a larger number of shallow cliffs as compared to *T. vaginalis* (Table S1).

3.3.6. Consensus models of the activity landscape: Consensus SAS maps

The activity landscape depends on the molecular representation.^{26,27} However, the consensus pairs found in several regions of the landscape suggests the possibility to derive, at least approximately, consensus models of the activity landscape. To this end, we employed in this work the principles of data fusion³⁸ producing Consensus SAS maps. For each pair of compounds, we calculated the mean and standard deviation of radial, MACCS, TGD and piDAPH3 similarity values (four selected orthogonal fingerprints) as detailed in Section 2. Figures 3E and 4E show the consensus SAS maps for *T. vaginalis* and *G. intestinalis*, respectively. In these figures, data points are colored by the most active compound in the pair using the same color scheme as in the SAS maps. Consensus SAS maps showing data points sized by standard deviation are in Figure S5 of the Supporting information. Figures 3E and 4E show the position of representative consensus pairs previously identified by radial, MACCS, TGD and piDAPH3 fingerprints (Figs. 3A–D and 4A–D). The mean fingerprint similarity values and standard deviation is provided in Table 4. In general, the pair of compounds in the consensus SAS map in Figure 3E occupies a relatively similar position (regions I–IV) in the SAS maps obtained separately with the radial, MACCS, TGD and piDAPH3 fingerprints (Fig. 3A–D). Similar results were obtained comparing the consensus SAS map of *G. intestinalis* (Fig. 4E) with the SAS maps obtained independently with the four fingerprints (Fig. 4A–D). Therefore, the consensus SAS maps effectively capture the information obtained with the different independent fingerprint representations and provide a good approximation of the overall activity landscape of the benzimidazoles tested against *T. vaginalis* and *G. intestinalis*. These results suggested that consensus SAS maps could provide valuable information for other data sets with other biological endpoints.⁴²

3.5. Dual-parasite SAR and consensus selectivity cliffs

We compared the activity of the 32 benzimidazoles against the two parasites. The difference of pIC₅₀ values is indicated in Table 1. Several compounds showed a similar activity (low Δ pIC₅₀) with *T. vaginalis* and *G. intestinalis*. For example, compounds **13**, **14**, **17–19**, **20**, **22–24**, **27** and **28** have a $|\Delta$ pIC₅₀| < 0.20. These results suggest that these non 2-methylcarbamates have a common mechanism of action against the two protozoan.⁴³ In addition, these results encourage the simultaneous lead optimization of compounds active against both *T. vaginalis* and *G. intestinalis*.

We also identified molecules with large potency difference against the two parasites. Compounds **1**, **6** and **12** are selective for *G. intestinalis*, whereas **4** and **9** are selective for *T. vaginalis* with more than one log unit in potency difference, respectively (Table

1). Interestingly, the pair of compounds **9_12** has a high structural similarity; the only difference is a CF₂ group (this difference was captured by all similarity methods, Table 4). However, the selectivity is quite different. This is an example of a selectivity cliff where a small change in the structure has a major impact in the selectivity.²⁵ A second example of a selectivity cliff was the pair **1_9** indicating that introducing bromine atoms in 2-(trifluoromethyl)benzimidazole has an opposite effect in the selectivity profile (Table 1). The pair **8_9** also has high structural similarity (Table 4; the difference is two bromine atoms at positions 4 and 7) but produces an opposite change in the activity against the two parasites. However, the effect for the pair **8_9** is less dramatic than for the pairs **9_12** and **1_9**. To note, **1_9** and **9_12** are deep cliffs while **8_9** is a shallow cliff in the landscape of *T. vaginalis*. However, the same pairs are in a smooth region of the landscape of *G. intestinalis*.

4. Conclusions and perspectives

We report a systematic characterization of the SAR of 32 (non 2-methylcarbamate) benzimidazoles with activity against *T. vaginalis* and *G. intestinalis*. The analysis was based on pairwise comparisons of the activity similarity and molecular similarity using different molecular representations. We found that radial, MACCS keys, TGD and piDAPH3 fingerprints showed low correlation (<0.80) and similar (approximately normal) distributions. These fingerprints, along with molecular properties, were used to characterize the activity landscape for each parasite. To note, the purpose of using multiple molecular representations was not to identify the 'best' representation but to identify the set of fingerprint representations that best identifies consensus data pairs. The landscape was portrayed using structure–activity similarity (SAS) maps which were quantitatively compared using the degree of consensus. The overall good consensus between structural representations allowed for the development of consensus models and consensus SAS maps to represent the activity landscape of *T. vaginalis* and *G. intestinalis*. For both parasites, several consensus pairs of compounds were identified in the smooth region of the landscape. Also a number of pairs were identified in the side chain hopping region. For *T. vaginalis*, we identified two deep consensus activity cliffs (**1_9** and **9_12**). It is anticipated that these compounds will be apparent 'outliers' in traditional computational models such as QSAR. In contrast, for *G. intestinalis*, we identified apparent cliffs and shallow cliffs. In conclusion, a heterogeneous SAR was found for both parasites. Comparison of the compounds' selectivity for each parasite revealed that several compounds are active against *T. vaginalis* and *G. intestinalis* showing similar SAR. We concluded that these molecules may have similar mechanism of action in both parasites and encourage simultaneous lead optimization efforts against both organisms. However, we also detected molecules with opposite selectivity profile and consensus selectivity cliffs.

Despite the fact that the data set of molecules studied in this work share the same core scaffold, whole-molecule fingerprint-based similarity methods were able to study the activity landscape, derive a consensus model of the landscape and, in particular, detect activity cliffs. Radial fingerprints were able to distinguish the molecules in great detail and differentiate positional isomers.

We want to emphasize that the present study is a systematic description of the activity landscape of a data set of 32 benzimidazole derivatives. This systematic study of the SAR will be expanded to larger data sets including compounds currently synthesized and tested. A second major perspective of this work is to explore the predictive capabilities of the activity landscape models to anticipate the SAR of new molecules in a prospective manner. This is an area of intense research in our and other

research groups. The systematic approach presented here to develop consensus models of the activity landscape of benzimidazole analogues against *T. vaginalis* and *G. intestinalis* is general. The approach can be applied to other larger data sets with other biological end points.

Acknowledgements

The authors would like to express their sincere thanks to Dr. Gerald M. Maggiora, for insightful discussions. We are also very grateful to Karen Gottwald for proofreading the manuscript. This work was supported by the State of Florida, Executive Office of the Governor's Office of Tourism, Trade, and Economic Development. J.P-V is supported by CONACyT and Carso Health Institute PhD studentships. J.P-V also thanks the State of Florida for partial support of his research visit at TPIMS.

Supplementary data

Total number of shallow cliffs according to different molecular representations (Table S1); active regions in the SAS maps for *T. vaginalis* (Fig. S1); active regions in the SAS maps for *G. intestinalis* (Fig. S2); examples of binned SAS maps (Fig. S3); number of consensus pairs between different molecular representations (Fig. S4); consensus SAS maps showing data points sized by standard deviation (Fig. S5). Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2010.09.019.

References and notes

- World Health Organization. Global Prevalence and Incidence of Selected Curable Sexually Transmitted Infections: Overview and Estimates, Geneva, 2001.
- Upcroft, P.; Upcroft, J. A. *Clin. Microbiol. Rev.* **2001**, *14*, 150.
- Berkman, D. S.; Lescano, A. G.; Gilman, R. H.; Lopez, S.; Black, M. M. *Lancet* **2002**, *359*, 564.
- Navarrete-Vázquez, G.; Rojano-Vilchis, M. D.; Yepez-Mulia, L.; Melendez, V.; Gerena, L.; Hernandez-Campos, A.; Castillo, R.; Hernandez-Luis, F. *Eur. J. Med. Chem.* **2006**, *41*, 135.
- Navarrete-Vázquez, G.; Yépez, L.; Hernández-Campos, A.; Tapia, A.; Hernández-Luis, F.; Cedillo, R.; González, J.; Martínez-Fernández, A.; Martínez-Grueiro, M.; Castillo, R. *Bioorg. Med. Chem.* **2003**, *11*, 4615.
- Andrzejewska, M.; Yépez-Mulia, L.; Cedillo-Rivera, R.; Tapia, A.; Vilpo, L.; Vilpo, J.; Kazimierczuk, Z. *Eur. J. Med. Chem.* **2002**, *37*, 973.
- Valdez-Padilla, D.; Rodríguez-Morales, S.; Hernández-Campos, A.; Hernández-Luis, F.; Yépez-Mulia, L.; Tapia-Contreras, A.; Castillo, R. *Bioorg. Med. Chem.* **2009**, *17*, 1724.
- Ooms, F. *Curr. Med. Chem.* **2000**, *7*, 141.
- Medina-Franco, J. L.; Lopez-Vallejo, F.; Castillo, R. *Educ. Quím.* **2006**, *17*, 452.
- Kubinyi, H. *Drug Discovery Today* **1997**, *2*, 457.
- Kubinyi, H. *Drug Discovery Today* **1997**, *2*, 538.
- Scior, T.; Medina-Franco, J. L.; Do, Q. T.; Martínez-Mayorga, K.; Yunes Rojas, J. A.; Bernard, P. *Curr. Med. Chem.* **2009**, *16*, 4297.
- Klebe, G. J. *Mol. Med.* **2000**, *78*, 269.
- Peltason, L.; Bajorath, J. *Chem. Biol.* **2007**, *14*, 489.
- Maggiora, G. M. *J. Chem. Inf. Model.* **2006**, *46*, 1535.
- Guha, R.; VanDrie, J. H. *J. Chem. Inf. Model.* **2008**, *48*, 646.
- Bajorath, J.; Peltason, L.; Wawer, M.; Guha, R.; Lajiness, M. S.; Van Drie, J. H. *Drug Discovery Today* **2009**, *14*, 698.
- Schneider, G.; Neidhart, W.; Giller, T.; Schmid, G. *Angew. Chem., Int. Ed.* **1999**, *38*, 2894.
- Eckert, H.; Bajorath, J. *Drug Discovery Today* **2007**, *12*, 225.
- Medina-Franco, J. L.; Martínez-Mayorga, K.; Giulianotti, M. A.; Houghten, R. A.; Pinilla, C. *Curr. Comput.-Aided Drug Des.* **2008**, *4*, 322.
- Peltason, L.; Bajorath, J. *J. Med. Chem.* **2007**, *50*, 5571.
- Guha, R.; Van Drie, J. H. *J. Chem. Inf. Model.* **2008**, *48*, 1716.
- Shanmugasundaram, V.; Maggiora, G. M. 222nd ACS National Meeting, Chicago, IL, United States; American Chemical Society, Washington, D. C.: Chicago, IL, United States, 2001.
- Wawer, M.; Peltason, L.; Weskamp, N.; Teckentrup, A.; Bajorath, J. *J. Med. Chem.* **2008**, *51*, 6075.
- Peltason, L.; Hu, Y.; Bajorath, J. *ChemMedChem* **2009**, *4*, 1864.
- Medina-Franco, J. L.; Martínez-Mayorga, K.; Bender, A.; Marín, R. M.; Giulianotti, M. A.; Pinilla, C.; Houghten, R. A. *J. Chem. Inf. Model.* **2009**, *49*, 477.
- Peltason, L.; Iyer, P.; Bajorath, J. *J. Chem. Inf. Model.* **2010**, *50*, 1021.
- Navarrete-Vázquez, G.; Cedillo, R.; Hernández-Campos, A.; Yépez, L.; Hernández-Luis, F.; Valdez, J.; Morales, R.; Cortés, R.; Hernández, M.; Castillo, R. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 187.
- Valdez, J.; Cedillo, R.; Hernandez-Campos, A.; Yepez, L.; Hernandez-Luis, F.; Navarrete-Vazquez, G.; Tapia, A.; Cortes, R.; Hernandez, M.; Castillo, R. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2221.
- Jaccard, P. *Bull. Soc. Vaudoise Sci. Nat.* **1901**, *37*, 547.
- Willett, P.; Barnard, J. M.; Downs, G. M. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 983.
- Molecular Operating Environment (MOE), version 2009.10, Chemical Computing Group, Montreal, Quebec, Canada. <http://www.chemcomp.com> (accessed August, 2010).
- Canvas, version 1.3, Schrödinger, LLC, New York, NY, 2010.
- Sastry, M.; Lowrie, J. F.; Dixon, S. L.; Sherman, W. *J. Chem. Inf. Model.* **2010**, *50*, 771.
- Conformational analysis for molecules with rotatable bonds was conducted by means of systematic search with torsions of 60° using MMFF94 force field. Equilibrium geometry calculations were carried out for the minimum energy conformer using PM3 semiempirical method.
- Spartan, version 2002, Wavefunction, Irvine, CA. <http://www.wavefun.com> (accessed August, 2010).
- A similar term, 'H-hopping' was proposed by Prof. Jürgen Bajorath (personal communication).
- Willett, P. *Drug Discovery Today* **2006**, *11*, 1046.
- Rogers, D.; Hahn, M. *J. Chem. Inf. Model.* **2010**, *50*, 742.
- The only exception were regions II and III of the SAS map for *T. vaginalis* obtained with TGD; both regions have similar population.
- Johnson, M. A.; Maggiora, G. M. *Concepts and Applications of Molecular Similarity*; Wiley: New York, 1990.
- By analogy with the SAS maps, consensus SAS maps could be analyzed quantitatively dividing the maps into four quadrants by imposing thresholds for activity and molecular similarity.
- Related methyl 1*H*-benzimidazole-2-yl-carbamate derivatives are strong selective inhibitors of tubulin polymerization. However, it is required that the benzimidazole scaffold has a 2-methylcarbamate group and hydrogen at position 1. The mechanism of non 2-methylcarbamates remains to be determined.

Cite this: *Med. Chem. Commun.*, 2011, **2**, 44

www.rsc.org/medchemcomm

CONCISE ARTICLE

Structure–activity relationships of benzimidazole derivatives as antiparasitic agents: *Dual activity-difference (DAD) maps*[†]Jaime Pérez-Villanueva,^a Radleigh Santos,^b Alicia Hernández-Campos,^a Marc A. Giulianotti,^b Rafael Castillo^a and Jose L. Medina-Franco^{*b}

Received 13th September 2010, Accepted 18th October 2010

DOI: 10.1039/c0md00159g

Parasitic infections still remain a major health threat in developing countries. Herein we report a systematic characterization of the structure–activity relationships (SAR) of a comprehensive set of benzimidazole derivatives tested against *Trichomonas vaginalis* and *Giardia intestinalis*. The analysis was based on pairwise comparisons of the activity similarity and molecular similarity using different molecular representations. Overall, results encourage simultaneous lead optimization efforts for benzimidazole derivatives active against both protozoan. In order to explore the activity profile of the benzimidazoles against the two parasites, we developed the *dual activity-difference (DAD) map*. DAD map is a complementary approach to systematically characterize the SAR of compound data sets.

Parasitic diseases are still a major health problem in developing countries. Among the most common protozoa infections are giardiasis, caused by *Giardia intestinalis*, and trichomonosis, a genitourinary infection caused by *Trichomonas vaginalis*.^{1,2} Recently, antiprotozoal activity of benzimidazole derivatives has been reported against these two parasites including compounds in the low nanomolar range. We have reported a first study towards the systematic characterization of the structure–activity relationships (SAR) of this class of compounds. Using the emerging concept of *activity landscape*³ we explored the SAR of 32 benzimidazoles finding a heterogeneous SAR for both parasites.⁴

Expanding on our previous work, herein we report a systematic characterization of the activity landscape of a larger set with 55 benzimidazole derivatives tested against *T. vaginalis* and *G. intestinalis*. This is a comprehensive set of published benzimidazoles with activity against the two parasites.^{5–9} The extended set includes 14 compounds with an alkylthio substituent at position 2⁸ and nine 2-(trifluoromethyl)benzimidazoles,^{5–7} for which the SAR has not been characterized thoroughly. We also present

a general approach to systematically explore the selectivity profile of these compounds against the two protozoan.

The chemical structures of the benzimidazole derivatives analyzed in this work are schematically depicted in Fig. 1. Most of the compounds have been synthesized in our group. Table S1 in the ESI,[†] shows the complete chemical structure for all 55 compounds. Noteworthy, 23 molecules in the data set have two tautomeric forms yielding a total of 78 chemical representations. Although theoretical efforts have been conducted to suggest the bioactive tautomeric form of some benzimidazoles as antiparasitic agents,¹⁰ it remains to be proven which tautomeric form(s) of the compounds are the biologically relevant for giardicidal and trichomonocidal activity. Table S1 also lists the 50% inhibitory

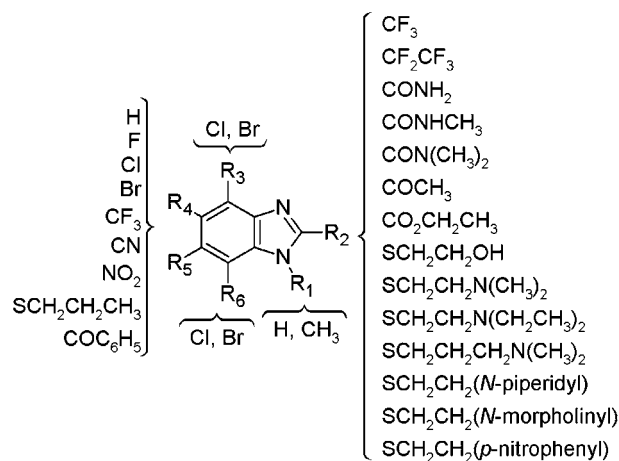


Fig. 1 General form of the benzimidazole derivatives studied in this work. The chemical structures of the 55 compounds, including the tautomeric forms, are shown in Table S1 in the ESI.[†]

^aFacultad de Química, Departamento de Farmacia, Universidad Nacional Autónoma de México, México, DF 04510, México

^bTorrey Pines Institute for Molecular Studies, 11350 SW Village Parkway, Port St. Lucie, FL 34987, USA. E-mail: jmedina@tpims.org; Fax: +1 772 345 3649; Tel: +1 772 345 4685

[†] Electronic supplementary information (ESI) available: Chemical structures and activity of benzimidazoles (Table S1); correlation matrix for the pairwise molecular similarities (Table S2); pairwise molecular similarities as cumulative distribution functions (Figure S1), SAS maps with different representations (Figure S2), correlation between the pIC₅₀ values for *T. vaginalis* and *G. intestinalis* (Figure S3), DAD map with data pairs color-coded using a continuous gray scale (Figure S4), SAS maps with data pairs color-coded using a continuous gray scale (Figure S5). See DOI: 10.1039/c0md00159g

concentration (IC₅₀) in *in vitro* susceptibility assays for each parasite as pIC₅₀ (−logIC₅₀). The activity of the 55 compounds was obtained by the same group under similar conditions. All compounds are non 2-methylcarbamates and their mechanism of action remains unknown.^{11,12}

The characterization of the antiprotozoal activity landscape for the benzimidazole derivatives was based on pairwise comparisons of the activity similarity and molecular similarity. Since chemical space^{13,14} and the activity landscape may vary with the structural representation,^{15,16} we considered an initial set of ten fingerprint representations to measure molecular similarity, namely radial, dendritic, MOLPRINT 2D, atom pairs, MACCS (166-bits), TDG, GpiDAPH3, piDAPH3, 3-, and 4-point pharmacophore fingerprints (see Methods). Despite the fact that the data set of molecules studied in this work share the same core scaffold, whole molecule fingerprint-based similarity methods were able to capture the activity landscape. Molecular similarity was computed with the Tanimoto coefficient. Figure S1 in the ESI summarizes the distribution of the 3,003 pairwise similarities for the 55 benzimidazoles (78 structural representations, considering the tautomeric forms) calculated with the ten fingerprints.† Table S2 in the ESI shows the correlation between the ten fingerprints.† Based on the distribution of similarity values and the relationships between the fingerprints, we selected dendritic, MACCS keys and piDAPH3 to characterize the activity landscape. Of note, the three fingerprints showed low correlation (<0.75) and dendritic fingerprint was able to differentiate between tautomeric forms.

Activity landscapes were portrayed using the well-established Structure–Activity Similarity (SAS) maps.¹⁷ Briefly, in a SAS

map, the activity similarity is represented in the Y-axis and molecular similarity is plotted in the X-axis. A general form is presented in Table 1. Using thresholds for activity and molecular similarity, it is possible to define four regions in the SAS maps. Region I corresponds to pairs of molecules with high activity similarity and low molecular similarity and therefore are associated with regions of scaffold (or side chain) hopping. Points in region II denote pairs of molecules with high molecular similarity and high activity similarity. Thus, compounds in this region are in a smooth or continuous SAR landscape. Region IV identifies pairs of molecules that have high molecular similarity and low activity similarity and therefore correspond to *activity cliffs* or discontinuous SAR.¹⁸ Figure S2 in the ESI shows six SAS maps,† three for each parasite using the selected molecular representations. For each parasite, we identified several pairs of compounds that are located in similar regions of the SAS map for all three molecular representations, *e.g.*, *consensus pairs* (marked in black in Figure S2†).¹⁶ Among the consensus pairs, we identified pairs of compounds with the same molecules in the pair that are located in similar regions of the SAS maps for *T. vaginalis* and *G. intestinalis*, *e.g.*, *dual-target consensus pairs*. Table 1 summarizes the number of data pairs in different regions of the SAS maps for each parasite (see Methods for the thresholds used to define the four regions). Table 1 also summarizes the number of consensus pairs and dual-target consensus pairs. Noteworthy, we observed a large number of dual-target consensus pairs, in particular in the side chain hopping region of the landscape (1,104 (37%) data points in region I). This result suggests a common mechanism of action for several of the benzimidazoles in both parasites. Similar

Table 1 Distribution of data points in different regions of the SAS maps.^a

		I		II		III		IV	
Representation	Median similarity of actives ^b	Total ^c	Active pairs ^d	Total	Active pairs	Total	Active pairs	Total	Active pairs
<i>T. vaginalis</i>									
Dendritic	0.26	1815	621	868	287	290	192	30	12
MACCS	0.54	1598	571	1085	337	237	158	83	46
piDAPH3	0.69	1959	635	724	237	289	177	31	27
Consensus pairs ^e		1388	499	509	194	217	142	9	9
<i>G. intestinalis</i>									
Dendritic	0.24	1575	696	898	397	472	341	58	29
MACCS	0.55	1535	715	938	378	393	280	137	90
piDAPH3	0.67	1645	635	828	458	477	327	53	43
Consensus pairs		1244	556	548	294	354	246	17	9
Dual-target consensus pairs ^f		1104	300	491	160	122	70	4	4

^a Regions I–IV are defined by the median similarity of active compounds and 0.5 as the threshold for activity similarity. ^b Median similarity of compounds with pIC₅₀ > 7.3 (IC₅₀ < 50 nM). ^c Total number of data points (pair of compounds) in the region. ^d Number of data points with at least one active compound (pIC₅₀ > 7.3) in the pair. ^e Consensus pairs for the three molecular representations. ^f Consensus pairs in both parasites for the three molecular representations.

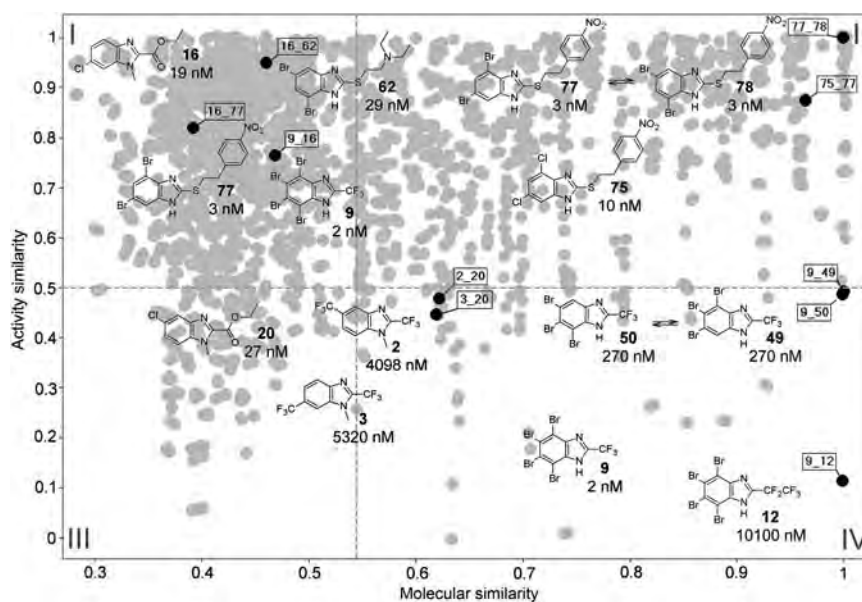


Fig. 2 Representative SAS map for *T. vaginalis* obtained with MACCS keys. Each data point indicates a pairwise comparison. Figure also shows the chemical structures and biological activity (IC_{50}) of selected consensus pairs in the three most representative regions of the SAS map.

conclusions were obtained with the smaller set of 32 compounds previously reported.⁴ We also detected four consensus activity cliffs common for both parasites (**2_16**, **2_20**, **3_16**, and **3_20**). Overall, a heterogeneous SAR was found for both protozoan.

Fig. 2 depicts a representative SAS map for *T. vaginalis*. The map shows the relationship between activity similarity and molecular similarity obtained with MACCS keys. The plot contains 3,003 data points that represent a pairwise comparison. Representative consensus pairs are shown along with their chemical structure and IC_{50} value against *T. vaginalis*. Compound pairs **9_16**, **16_62**, and **16_77**, in region I of the SAS map, have low structural similarity and high activity similarity (Fig. 2). These data points are examples of side chain hopping, *i.e.*, the substitution pattern around the benzimidazole scaffold is quite different; however, all compounds have similar activity. The compound pair **75_77**, in region II, has high structural similarity and high activity similarity and illustrates a smooth SAR. Data point **77_78**, also in region II, represents a tautomeric pair that is approximated to have the same activity. The pair **9_12**, in region IV, illustrates an activity cliff where a small change in the structure is associated with a significant change in the IC_{50} .¹⁸ Additional examples of activity cliffs illustrated in Fig. 2 are the pairs **2_20**, **3_20**, **9_49**, and **9_50**. Although the change in activity for these pairs is less dramatic than the change in IC_{50} for **9_12**, it is readily observed that small changes in the substitution pattern of the benzimidazole scaffold are associated with considerable changes in activity. All pairs of compounds discussed here have the same relative position in the SAS maps of *T. vaginalis* obtained with other structural representations (Figure S2†). Moreover, most of these pairs occupy the same relative position in the SAS maps of *G. intestinalis*. Exceptions are pairs **9_12**, **9_49**, and **9_50** that are in a smooth region of the SAR for *G. intestinalis* (Figure S2†).

Having characterized the activity landscape of the benzimidazole derivatives for *T. vaginalis* and *G. intestinalis*, we

systematically compared the activity and SAR of the molecules against the two parasites. The difference of pIC_{50} values is indicated in Table S1.† Several compounds showed a similar activity (low ΔpIC_{50}) with *T. vaginalis* and *G. intestinalis*. We observed that most of the highly active compounds against *T. vaginalis* ($pIC_{50} > 7$, $IC_{50} < 100$ nM) also have high activity against *G. intestinalis*. These compounds have large substituents at the position 2 of the benzimidazole scaffold such as alkylthio and ethyl ester (**16**, **20**, **23**, **24**, **28**, **32**, **55**, **59–68**, **71–78**). Furthermore, for both parasites trifluoromethyl substitution at position 2 give compounds with activities that ranges from poor to moderate ($pIC_{50} = 4.53–6.97$) with the exception of compound **9** which has strong activity. Interestingly, the 1-methylbenzimidazole derivatives **13–32** have low differences in activity against two parasites in almost all cases ($|\Delta pIC_{50}| = 0–0.77$, median $|\Delta pIC_{50}| = 0.205$). For these compounds, chlorine atom substitution at position 6 is favorable for the activity for both parasites while substitution at position 5 is less favorable. Similar conclusions were obtained with the previous data of the 32 non 2-methylcarbamates.⁴ These results suggest that several of the new 2-(trifluoromethyl)benzimidazoles and compounds with an alkylthio substituent at position 2 also have a common mechanism of action against the two protozoan. However, we also identified molecules with large potency difference against the two parasites; there are nine compounds with more than one log unit in potency difference, and five compounds with more than 1.4 log units (Table S1†). None of the five compounds have an alkylthio substituent at position 2. Most of these compounds with large activity differences have bromine atoms at positions 4 to 7 or large substituents at positions 5 and 6.

The correlation between the pIC_{50} values for *T. vaginalis* and *G. intestinalis* is shown in Figure S3 in the ESI.† From this correlation it is straightforward to identify the five compounds discussed above with high potency differences. However, using just the correlation of pIC_{50} values or potency differences

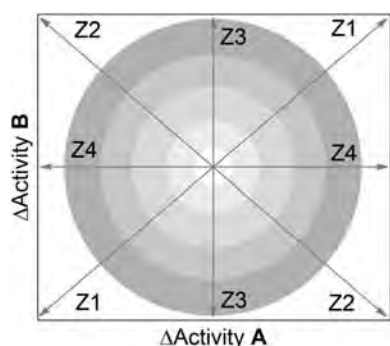


Fig. 3 General form of the dual activity-difference (DAD) map showing four major zones, labeled as Z1–Z4. Concentric rings denote different magnitudes in activity differences; outer rings indicate larger differences. Pairs of compounds along or nearby Z1, indicate that the activity of the two compounds in the pair against **A** and **B** increases or decreases in similar proportion. Pairs of compounds along or nearby Z2, indicate that the change in activity for the compounds in the pair is opposite for **A** and **B**. Zones 3 and 4 identify pairs of compounds with the same or similar activity for one target, but different activity for the other target. Data pairs located at the intersections of Z1–Z4 (inner circle) indicates that the pair of compounds has similar activity against **A** and **B**. See text for details.

(ΔpIC_{50}), it is not easy to systematically deduce the structural modifications associated with changes in potency against the two parasites. To this end, herein we propose the dual-activity difference (DAD) map. A general form of the map is in Fig. 3. This map is based on pairwise activity differences against the two parasites; the activity difference for one parasite is represented in the Y-axis and the activity difference for the second parasite is plotted in the X-axis.

A DAD map provides a visual and systematic characterization of the selectivity profile for a set of compounds screened with two parasites (or, in general, with two “targets”) **A** and **B**. Four general zones can be distinguished in Fig. 3, delineated by the lines Z1–Z4. Concentric rings in the figure denote different magnitudes in activity difference; outer rings indicate larger differences. Pairs of compounds along line Z1 or nearby this zone, indicate that the activity of the two compounds in the pair against **A** and **B** increases or decreases in similar proportion. In other words, the structural changes for the molecules in the pair along Z1 (either a small or a large structural change) have a similar impact in the activity against the two targets (either an increase or decrease in activity). Therefore, zone Z1 is associated with similar SAR of the compounds for both targets. Pairs of compounds along or nearby Z2, indicate that the change in activity for the compounds in the pair is opposite for **A** and **B**. Thus, the structural change in the pair of compounds along Z2 is associated with an *inverse SAR*, *i.e.*, increases the activity for one target but decreases the activity for the other target. Data points along or nearby Z3 and Z4, denote pairs of molecules with the same or similar activity for one target (**A** or **B**, respectively), but different activity for the other target (**B** or **A**, respectively). Data pairs located exactly or close to the intersections of lines Z1–Z4 (inner circle) indicates that the pair of compounds has identical or similar activity against **A** and **B**. In other words, structural changes (either small or large changes) in the pair of compounds

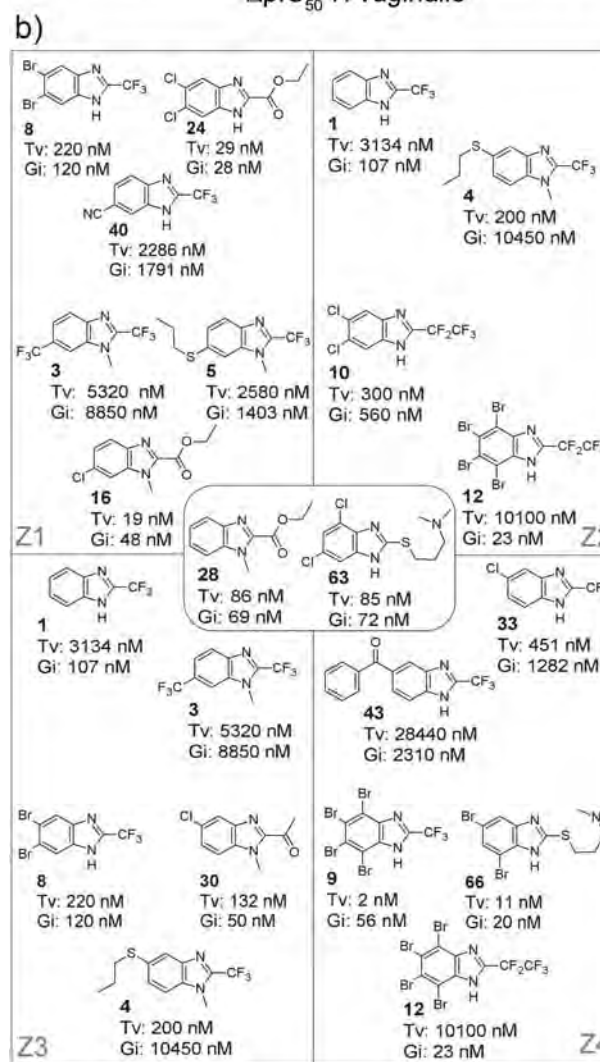
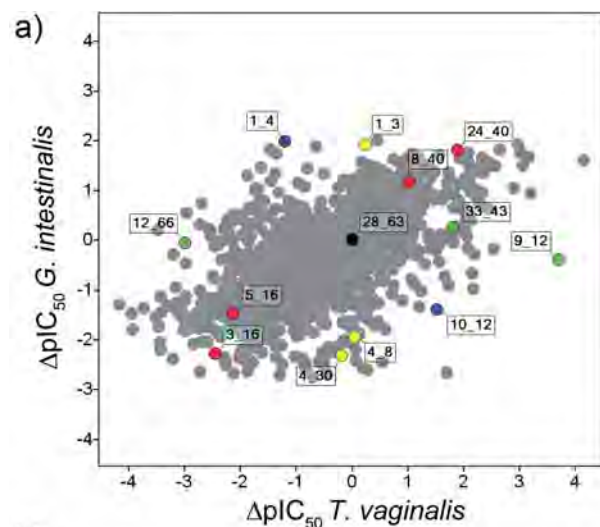


Fig. 4 (a) DAD map for the benzimidazole derivatives tested with *T. vaginalis* and *G. intestinalis*. Selected pairs are color-coded by zone in the map (see also Fig. 3): red, Z1; blue, Z2; yellow, Z3; and green, Z4. The black data point is approximately at the center of the map. (b) Chemical structures of selected pairs in panel A along with their IC_{50} values for the two parasites.

in this region have no or little impact, in the activity against the two targets. Noteworthy, molecular similarity is not represented directly in Fig. 3. However, molecular similarity information can be added to the DAD map (see below).

Fig. 4A shows the DAD map for the benzimidazole derivatives tested with *T. vaginalis* and *G. intestinalis*. The plot contains 3,003 data points that represent all (non-duplicate) pairwise comparisons (see Methods). Selected data points are colored by the zone they are associated with (see also Fig. 3); red points are close to Z1, blue points are close to Z2, yellow points are close to Z3, and green points are close to Z4. The black data point is approximately at the center of the map *e.g.*, at the intersection of the four zones. Fig. 4B shows the chemical structures of the selected pairs in Fig. 4A along with their IC₅₀ values for *T. vaginalis* and *G. intestinalis*. The pair **28_63**, located towards the center of the map, exemplifies a pair of molecules where the structural changes do not modify significantly the activity (either for *T. vaginalis* (IC₅₀ = 86 vs. 85 nM) or for *G. intestinalis* (IC₅₀ = 69 vs. 72 nM)). Of note, **28_63** is also a good example of side chain hopping for both parasites, *i.e.*, the substitution pattern around the benzimidazole scaffold is quite different; however, the activity is very similar (Fig. 4B).

Compound pairs **3_16**, **5_16**, **8_40**, and **24_40** are examples of data points close to Z1. The structural changes in each of these molecular pairs are associated with a similar change in activity difference for *T. vaginalis* and *G. intestinalis*. For example, for the pair **3_16** the change at R₂ of a trifluoromethyl group (**3**) to ethyl ester (**16**), and the change at R₅ of a trifluoromethyl group (**3**) to a chlorine atom (**16**), are associated with an increase in activity for both parasites (Fig. 4B). Similar conclusions can be obtained for other pairs in region Z1. Also note, the activity difference (ΔpIC_{50}) for both parasites is larger in pairs **3_16** and **24_40** (located in outer regions of the map) as compared with the activity differences observed in pairs **5_16** and **8_40** (located towards the center of the map).

Compound pairs **1_4** and **10_12** are examples of data points close to the Z2 line (Fig. 4A). The structural change in each pair is associated with an inverse effect in activity. For example, comparing pair **1_4**, *N*-methylation and substitution at R₄ for a propylthio group (**4**) is associated with a significant increase in activity for *T. vaginalis*. However, the same structural changes are associated with a significant decrease in activity for *G. intestinalis* (Fig. 4B). Of note, **1** is selective for *G. intestinalis* whereas **4** is selective for *T. vaginalis*.

Data points **1_3**, **4_8**, and **4_30** are examples of pairs of compounds close to Z3 (Fig. 4A). In these points, the activity against *T. vaginalis* is similar whereas the activity against *G. intestinalis* changes (Fig. 4B). In contrast, compound pairs **9_12**, **12_66**, and **33_43**, close to Z4, exemplify data points with similar activity against *G. intestinalis* but different activity against *T. vaginalis*. As discussed above, data points in the outer regions of the map have larger differences in activity (see concentric rings in Fig. 3). It is worth mentioning the pair **9_12**, with similar activity for *G. intestinalis* but 3.7 log units in potency difference for *T. vaginalis* (Fig. 4A and 4B). Of note, **12** is selective for *G. intestinalis* whereas **33** is selective for *T. vaginalis*.

Molecular similarity information is not represented in the general form of the DAD map. Thus, it is not possible to deduce from this map if, for a given pair, the activity difference is

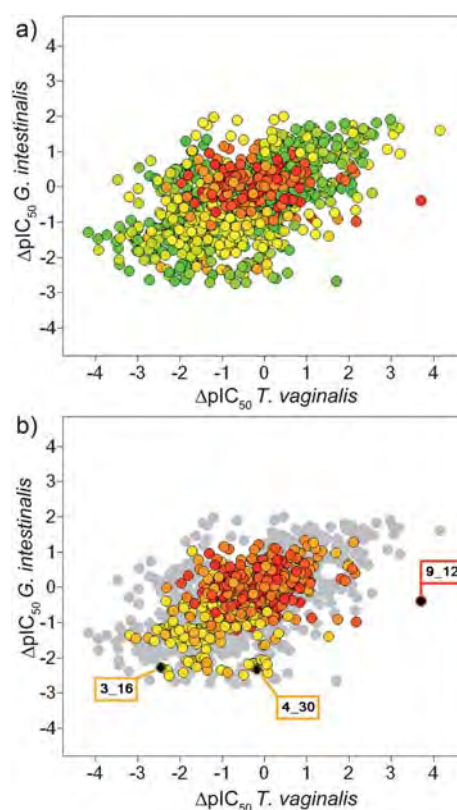


Fig. 5 (a) DAD map with data pairs color-coded by dendritic fingerprint similarity using a continuous scale from green (less similar) to red (more similar). (b) Map showing in color 925 data pairs with high similarity. Selected pairs are marked in black and labeled with the compound numbers. The chemical structures of compounds in these pairs are shown in Fig. 4. See text for details. The figure using a continuous scale from white to black is shown in Figure S4 in the ESI.†

associated with a small or large molecular change. However, since the DAD map is based on pairwise comparisons, it is straightforward to represent pairwise molecular similarity into this plot. Fig. 5A shows the DAD map of the benzimidazole derivatives distinguishing the data points by dendritic fingerprint similarity. A continuous color scale is used from green (less similar; similarity = 0.016) to red (more similar; similarity = 0.92). Fig. 5B shows in color 925 data pairs with high molecular similarity (similarity higher than the median similarity of the active compounds, see Table 1). Similar maps were obtained using MACCS keys and piDAPH3 fingerprints (data not shown). In general, we observed that compounds with high structural similarity tend to be located towards the center of the DAD map (Fig. 5B). This means that several pairs of benzimidazoles with high structural similarity have low activity differences for both parasites. This result further supports the conclusion that several benzimidazole derivatives have a similar and smooth SAR against *T. vaginalis* and *G. intestinalis*. However, we also detected pairs of structurally similar molecules with larger potency differences against the two parasites (a prominent example is the pair **9_12**).

Mapping molecular similarity into the DAD plot clearly reveals activity differences associated with small (or large) differences in structure. Three representative examples are illustrated in Fig. 5B. Data pairs **3_16** and **4_30** in regions Z1 and

Z3, respectively, have high molecular similarity as measured with dendritic fingerprint. The pair **9_12** in region Z4 also has high molecular similarity and it is associated with a large change in activity difference for *T. vaginalis*. However, the activity difference with *G. intestinalis* is small. Therefore, it is possible to deduce from the DAD map, colored by molecular similarity that **9_12** is an activity cliff in the landscape of *T. vaginalis* but it is in the smooth region of the SAR for *G. intestinalis*.

In conclusion, a systematic characterization of the activity landscape of a comprehensive data set of benzimidazole derivatives screened against *T. vaginalis* and *G. intestinalis* revealed a heterogeneous SAR in both cases. Results supported our previous conclusions suggesting a common mechanism of action for several compounds against the two parasites, including compounds from the new set of 2-(trifluoromethyl)benzimidazoles and compounds with an alkylthio substituent at position 2. These results further encourage simultaneous lead optimization efforts for benzimidazole derivatives active against both protozoan. We also identified some compounds selective for each parasite and confirmed the presence of consensus activity cliffs. As part of the characterization of the SAR against both parasites, we developed the DAD map. This map allows for one to systematically explore differences in antiprotozoal activity for the benzimidazoles in order to identify structural changes associated with activity differences. Since molecular similarity information can be easily represented in the plot, the DAD map is a complementary tool used to systematically explore the SAR of the benzimidazoles. The predictive ability of the SAS and DAD maps in prospective studies remains to be explored.

Methods

The chemical structures and biological activity of the 55 benzimidazole derivatives were taken from the literature. Pairwise structural similarities were computed with the Tanimoto coefficient.¹⁹ 2D similarity values were calculated with the Molecular Operating Environment (MOE)²⁰ MACCS keys (166 bits), graph-based three point pharmacophores (GpiDAPH3), and typed graph distance (TGD). We also used the following 2D (32-bit) fingerprints as implemented in Canvas:^{21,22} radial, dendritic, atom pairs, and MOLPRINT 2D. 3D similarity values were calculated with the MOE pharmacophore atom triangle (piDAPH3) and Canvas 3- and 4-point pharmacophores fingerprints. To compute 3D similarities, a single low-energy conformation was considered for each molecule as previously reported.⁴ For each pair of molecules, the activity similarity for *T. vaginalis* and *G. intestinalis* was measured with the expression: $AS_{i,j} = 1 - \frac{|A_i - A_j|}{\max - \min}$, where A_i and A_j are the pIC₅₀ values of the *i*th and *j*th molecules, and max-min indicate the range of activities in the data set. SAS maps were constructed by plotting the activity similarity against the structural similarity for each pair of compounds. In order to quantitatively characterize the SAS maps, each map was partitioned in four regions by using 0.5 as threshold for activity similarity. The median similarity of the most active compounds in the data set (pIC₅₀ > 7.3) was used as threshold for molecular similarity.¹⁶ DAD maps were generated by plotting pairwise pIC₅₀ differences for both parasites. Duplicate comparisons were not considered. For example, compounds pairs **1_2** and **2_1** would

have the same coordinates in the DAD map, with opposite sign. In this example we just considered one pair, **1_2**.

Acknowledgements

The authors thank Dr Gerald M. Maggiora, for enlightening discussions. This work was supported by the State of Florida, Executive Office of the Governor's Office of Tourism, Trade, and Economic Development. J.P-V is supported by CONACyT and Carso Health Institute PhD studentships. J.P-V also thanks the State of Florida for partial support of his research visit at TPIMS.

References and notes

- 1 P. Upcroft and J. A. Upcroft, *Clin. Microbiol. Rev.*, 2001, **14**, 150–164.
- 2 D. S. Berkman, A. G. Lescano, R. H. Gilman, S. Lopez and M. M. Black, *The Lancet*, 2002, **359**, 564–571.
- 3 J. Bajorath, L. Peltason, M. Wawer, R. Guha, M. S. Lajiness and J. H. Van Drie, *Drug Discovery Today*, 2009, **14**, 698–705.
- 4 J. Pérez-Villanueva, R. Santos, A. Hernández-Campos, M. A. Giulianotti, R. Castillo and J. L. Medina-Franco, *Bioorg. Med. Chem.*, 2010, **18**, 7380–7391.
- 5 G. Navarrete-Vázquez, L. Yépez, A. Hernández-Campos, A. Tapia, F. Hernández-Luis, R. Cedillo, J. González, A. Martínez-Fernández, M. Martínez-Grueiro and R. Castillo, *Bioorg. Med. Chem.*, 2003, **11**, 4615–4622.
- 6 G. Navarrete-Vázquez, M. D. Rojano-Vilchis, L. Yépez-Mulia, V. Meléndez, L. Gerena, A. Hernández-Campos, R. Castillo and F. Hernández-Luis, *Eur. J. Med. Chem.*, 2006, **41**, 135–141.
- 7 M. Andrzejewska, L. Yépez-Mulia, R. Cedillo-Rivera, A. Tapia, L. Vilpo, J. Vilpo and Z. Kazimierzczuk, *Eur. J. Med. Chem.*, 2002, **37**, 973–978.
- 8 M. Andrzejewska, L. Yépez-Mulia, A. Tapia, R. Cedillo-Rivera, A. E. Laudy, B. J. Starościk and Z. Kazimierzczuk, *Eur. J. Pharm. Sci.*, 2004, **21**, 323–329.
- 9 D. Valdez-Padilla, S. Rodríguez-Morales, A. Hernández-Campos, F. Hernández-Luis, L. Yépez-Mulia, A. Tapia-Contreras and R. Castillo, *Bioorg. Med. Chem.*, 2009, **17**, 1724–1730.
- 10 F. López-Vallejo, J. L. Medina-Franco, A. Hernández-Campos, S. Rodríguez-Morales, L. Yépez, R. Cedillo and R. Castillo, *Bioorg. Med. Chem.*, 2007, **15**, 1117–1126.
- 11 G. Navarrete-Vázquez, R. Cedillo, A. Hernández-Campos, L. Yépez, F. Hernández-Luis, J. Valdez, R. Morales, R. Cortés, M. Hernández and R. Castillo, *Bioorg. Med. Chem. Lett.*, 2001, **11**, 187–190.
- 12 J. Valdez, R. Cedillo, A. Hernández-Campos, L. Yépez, F. Hernández-Luis, G. Navarrete-Vázquez, A. Tapia, R. Cortés, M. Hernández and R. Castillo, *Bioorg. Med. Chem. Lett.*, 2002, **12**, 2221–2224.
- 13 J. L. Medina-Franco, K. Martínez-Mayorga, M. A. Giulianotti, R. A. Houghten and C. Pinilla, *Curr. Comput.-Aided Drug Des.*, 2008, **4**, 322–333.
- 14 J.-L. Reymond, R. van Deursen, L. C. Blum and L. Ruddigkeit, *Med. Chem. Commun.*, 2010, **1**, 30–38.
- 15 L. Peltason, P. Iyer and J. Bajorath, *J. Chem. Inf. Model.*, 2010, **50**, 1021–1033.
- 16 J. L. Medina-Franco, K. Martínez-Mayorga, A. Bender, R. M. Marín, M. A. Giulianotti, C. Pinilla and R. A. Houghten, *J. Chem. Inf. Model.*, 2009, **49**, 477–491.
- 17 V. Shanmugasundaram and G. M. Maggiora, 222nd ACS National Meeting, Chicago, IL, United States, 2001. American Chemical Society, Washington, D.C., Meeting abstract CINF-032.
- 18 G. M. Maggiora, *J. Chem. Inf. Model.*, 2006, **46**, 1535–1535.
- 19 P. Willett, J. M. Barnard and G. M. Downs, *J. Chem. Inf. Comput. Sci.*, 1998, **38**, 983–996.
- 20 *Molecular Operating Environment (MOE), version 2009.10*, Chemical Computing Group Inc., Montreal, Quebec, Canada. Available at <http://www.chemcomp.com> (Accessed September, 2010).
- 21 M. Sastry, J. F. Lowrie, S. L. Dixon and W. Sherman, *J. Chem. Inf. Model.*, 2010, **50**, 771–784.
- 22 *Canvas, version 1.3*, Schrödinger, LLC, New York, NY, 2010.



Original article

Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) of some benzimidazole derivatives with trichomonocidal activity

Jaime Pérez-Villanueva^a, José L. Medina-Franco^b, Thomas R. Caulfield^b, Alicia Hernández-Campos^a, Francisco Hernández-Luis^a, Lilián Yépez-Mulia^c, Rafael Castillo^{a,*}

^a Facultad de Química, Departamento de Farmacia, UNAM, México, DF 04510, Mexico

^b Torrey Pines Institute for Molecular Studies, Port St. Lucie, FL 34987, USA

^c Unidad de Investigación Médica en Enfermedades Infecciosas y Parasitarias, IMSS, México, DF 06720, Mexico

ARTICLE INFO

Article history:

Received 5 November 2010

Received in revised form

6 May 2011

Accepted 7 May 2011

Available online 13 May 2011

Keywords:

Benzimidazole

3D-QSAR

CoMFA

CoMSIA

Structure–activity relationships

Trichomonas vaginalis

ABSTRACT

Trichomonosis is a common sexually transmitted infectious disease linked to reproductive health complications. Recently, the benzimidazole nucleus has emerged as a promising scaffold to develop new trichomonocidal agents. Despite the fact that large amounts of experimental data have been accumulated over the past eight years, no quantitative studies have yet been reported on this class of compounds. In our effort to develop new antiparasitic benzimidazole derivatives, we report in this paper CoMFA and CoMSIA studies with an initial set of 70 benzimidazole derivatives with trichomonocidal activity. Four CoMFA models and eight CoMSIA models were generated; ten of these models had values of $r^2 > 0.6$ and $q^2 > 0.5$. The best CoMFA model had $r^2 = 0.936$ and $q^2 = 0.634$, and the best CoMSIA model had $r^2 = 0.858$ and $q^2 = 0.642$. These models were generated by using two conformer selection methodologies (minimum energy conformations and 3D similarity), and three charge types (Mulliken, Gasteiger-Hückel and electrostatic potential atomic charges). The putative active tautomers of 1*H*-benzimidazole derivatives were selected using 3D-QSAR calculations. All models were validated via an external test set with 13 molecules. The best models satisfied additional validation criteria. The contour maps generated show the most important features that a benzimidazole derivative should have for trichomonocidal activity; they also, suggest that substituents at the 2- and 6-positions are important in the generation of derivatives with strong activity.

© 2011 Elsevier Masson SAS. All rights reserved.

1. Introduction

Trichomonosis is a sexually transmitted disease caused by the protozoan *Trichomonas vaginalis*. According to the World Health Organization (WHO), more than 170 million new cases are estimated annually worldwide [1]. In women, infection can cause profound inflammation of the genital tract which has been associated with preterm labor, low-birth weight, sterility, cervical cancer and a predisposition to HIV infection [2,3]. In males, *T. vaginalis* infection is generally asymptomatic; however, in some cases it causes urethritis and prostatitis [2,3]. Metronidazole is the drug of choice against this infection; never the less, its side effects and

the development of resistant strains have stimulated our interest to search for new therapeutic agents [3,4].

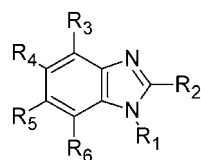
Compounds with the benzimidazole nucleus are commonly encountered in drugs having a broad diversity of pharmacological activities: anthelmintic, gastroprotective, antitumoral, antiviral, antifungal, and antibacterial [5]. As part of our efforts to find new benzimidazole derivatives as antiparasitic agents, a large number of this type of compounds against *T. vaginalis*, *Entamoeba histolytica* and *Giardia intestinalis* have previously been synthesized and tested [6–10]. The trichomonocidal benzimidazole derivatives that have been synthesized, together with that reported by another group [11,12], make up an important database in order to study the structure–activity relationships and propose new potentially active compounds using computer-aided drug design [13–16].

Comparative Molecular Field Analysis (CoMFA) [17,18] and Comparative Molecular Similarity Indices Analysis (CoMSIA) [19] are two powerful methodologies designed to study quantitative

* Corresponding author. Tel.: +52 5 56 22 52 87; fax: +52 5 56 22 53 29.

E-mail address: rafaelc@servidor.unam.mx (R. Castillo).

Table 1
Chemical structure and trichomonocidal activity of benzimidazole derivatives used in this study.



	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	pIC ₅₀ ^a
1	H	CF ₃	H	H	H	H	5.50
2	H	CF ₃	H	Cl	H	H	6.35
3 ^b	H	CF ₃	H	F	H	H	5.50
4	H	CF ₃	H	CF ₃	H	H	6.63
5	H	CF ₃	H	CN	H	H	5.64
6	CH ₃	CF ₃	H	CF ₃	H	H	5.39
7	CH ₃	CF ₃	H	H	CF ₃	H	5.27
8	H	CF ₃	H	SCH ₂ CH ₂ CH ₃	H	H	6.46
9	CH ₃	CF ₃	H	SCH ₂ CH ₂ CH ₃	H	H	6.70
10	CH ₃	CF ₃	H	H	SCH ₂ CH ₂ CH ₃	H	5.59
11 ^b	H	CF ₃	H	COC ₆ H ₅	H	H	4.55
12	CH ₃	CF ₃	H	COC ₆ H ₅	H	H	4.53
13	CH ₃	CF ₃	H	H	COC ₆ H ₅	H	4.97
14	H	H	H	H	H	H	6.44
15	H	CH ₃	H	H	H	H	6.52
16	H	NH ₂	H	H	H	H	6.54
17	H	SH	H	H	H	H	6.71
18	H	H	H	Cl	H	H	6.69
19	H	CH ₃	H	Cl	H	H	6.79
20	H	NH ₂	H	Cl	H	H	6.71
21	H	SH	H	Cl	H	H	6.87
22 ^b	H	SCH ₃	H	Cl	H	H	7.03
23	H	NH ₂	H	Cl	Cl	H	6.98
24	H	CF ₃	H	Br	H	H	5.80
25	H	CF ₃	H	Br	Br	H	6.66
26	H	CF ₃	Br	H	Br	H	6.72
27 ^b	H	CF ₃	Br	Br	Br	H	6.57
28	H	CF ₃	Br	Br	Br	Br	8.70
29	H	C ₂ F ₅	H	Cl	Cl	H	6.52
30	H	CF ₃	H	NO ₂	NO ₂	H	6.24
31	H	C ₂ F ₅	Br	Br	Br	Br	5.00
32	H	SCH ₂ CH ₂ OH	Cl	H	Cl	H	6.16
33	H	SCH ₂ CH ₂ OH	Br	H	Br	H	6.83
34	H	SCH ₂ CH ₂ N(CH ₃) ₂	Cl	H	Cl	H	7.14
35	H	SCH ₂ CH ₂ N(CH ₃) ₂	Br	H	Br	H	6.51
36 ^b	H	SCH ₂ CH ₂ N(C ₂ H ₅) ₂	Cl	H	Cl	H	7.39
37	H	SCH ₂ CH ₂ N(C ₂ H ₅) ₂	Br	H	Br	H	7.53
38	H	SCH ₂ CH ₂ CH ₂ N(CH ₃) ₂	Cl	H	Cl	H	7.07
39	H	SCH ₂ CH ₂ CH ₂ N(CH ₃) ₂	Br	H	Br	H	7.99
40	H	SCH ₂ CH ₂ -(<i>N</i> -piperidyl)	Cl	H	Cl	H	7.14
41 ^b	H	SCH ₂ CH ₂ -(<i>N</i> -piperidyl)	Br	H	Br	H	7.69
42	H	SCH ₂ CH ₂ -(<i>N</i> -morpholinyl)	Cl	H	Cl	H	7.29
43	H	SCH ₂ CH ₂ -(<i>N</i> -morpholinyl)	Br	H	Br	H	7.23
44	H	SCH ₂ CH ₂ -(<i>p</i> -nitrophenyl)	Cl	H	Cl	H	7.96
45	H	SCH ₂ CH ₂ -(<i>p</i> -nitrophenyl)	Br	H	Br	H	8.48
46 ^b	CH ₃	CONH ₂	H	H	Cl	H	6.96
47	CH ₃	CONHCH ₃	H	H	Cl	H	6.98
48 ^b	CH ₃	CON(CH ₃) ₂	H	H	Cl	H	6.63
49	CH ₃	COOCH ₂ CH ₃	H	H	Cl	H	7.72
50	CH ₃	CONH ₂	H	Cl	H	H	6.73
51	CH ₃	CONHCH ₃	H	Cl	H	H	6.45
52	CH ₃	CON(CH ₃) ₂	H	Cl	H	H	6.68
53 ^b	CH ₃	COOCH ₂ CH ₃	H	Cl	H	H	7.57
54 ^b	CH ₃	CONH ₂	H	Cl	Cl	H	6.87
55	CH ₃	CONHCH ₃	H	Cl	Cl	H	6.65
56	CH ₃	CON(CH ₃) ₂	H	Cl	Cl	H	7.12
57	CH ₃	COOCH ₂ CH ₃	H	Cl	Cl	H	7.53
58	CH ₃	CONH ₂	H	H	H	H	6.78
59 ^b	CH ₃	CONHCH ₃	H	H	H	H	6.98
60	CH ₃	CON(CH ₃) ₂	H	H	H	H	6.37
61	CH ₃	COOCH ₂ CH ₃	H	H	H	H	7.07
62	CH ₃	NHCOOCH ₃	H	H	H	H	5.88
63	CH ₃	NHCOOCH ₃	H	Cl	H	H	6.36
64	CH ₃	NHCOOCH ₃	H	H	Cl	H	6.46

Table 1 (continued)

	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	pIC ₅₀ ^a
65 ^b	CH ₃	NHCOOCH ₃	H	Cl	Cl	H	6.27
66	H	SCH ₃	H	CONH ₂	H	H	6.67
67	H	SCH ₃	H	CONHCH ₃	H	H	6.79
68	H	SCH ₃	H	CON(CH ₃) ₂	H	H	7.54
69 ^b	H	SCH ₃	H	CONHCH ₂ CH ₃	H	H	7.26
70	H	SCH ₃	H	CON(CH ₂ CH ₃) ₂	H	H	6.05

^a Activity data are expressed in pIC₅₀ = -log IC₅₀.

^b Compounds in test set.

structure–activity relationships (QSAR). In previous studies, the CoMFA methodology has been successful when applied in the study of benzimidazole derivatives as antitumorals [20], antivirals [21], 5HT₄ agonists [22] and recently our research group reported studies against *E. histolytica* [23]. In this paper, for the first time CoMFA and CoMSIA studies with a database of benzimidazole derivatives active against *T. vaginalis* are reported. The QSAR models show high quantitative correlations with good predictive abilities, highlighting the most important features in the design of new benzimidazole derivatives with trichomonocidal activity.

2. Materials and methods

2.1. Database and biological activity

The data set used in this study contained 70 benzimidazole derivatives which have been previously published (Table 1). All compounds have been synthesized by our group [8–10], except **24–45**, reported by Andrzejewska et al. [11,12]. All trichomonocidal activities were experimentally determined in the same laboratory using the *T. vaginalis* GT3 strain [8–12]. IC₅₀ values were converted to pIC₅₀ = -log IC₅₀, using the activity in mol/L.

2.2. Training and test set selection

The initial set of 70 compounds was divided into a training set with 57 (81%) molecules and a test set with 13 (19%) molecules. The activities of the training set ranges from 2 nM (pIC₅₀ = 8.70) to 29,512 nM (pIC₅₀ = 4.53); median pIC₅₀ = 6.70. The activities of the test set ranges from 20 nM (pIC₅₀ = 7.69) to 28,440 nM (pIC₅₀ = 4.55); median pIC₅₀ = 6.98. Worthy of mention is the fact that the structural diversity and activity range of the test set are comparable with the training set [18,24,25].

2.3. Conformer generation

Molecular modeling calculations were completed using the Spartan'02 software package and performed on a Dell Pentium IV PC running Red Hat Enterprise Linux 4 operative system [26]. Conformational analysis was conducted by a systematic search with increments of 60° for each rotatable bond with the MMFF94 force field; the conformational search was restricted to conformers with best energies within a 10 kcal/mol range. The conformers obtained and the molecules with no rotatable bonds were further optimized using the PM3 semiempirical method. Atomic charges (Mulliken and electrostatics) [27,28] were calculated with PM3 implemented in Spartan'02 software [29]. Gasteiger–Hückel atomic charges [30] were calculated with Sybyl 7.3 software [31].

Minimum energy conformers were selected. 3D similarity conformer selection was carried out as an alternate approach. Molecular overlap was conducted with the Rapid Overlay of Chemical Structures (ROCS) approach [32], using **9**, **10** and **45** as query structures on its minimum energy conformations. The combo

score, comprised of the Tanimoto shape index and color score, was used to select conformers. The Tanimoto index was calculated with equation (1),

$$T_{A,B} = \frac{O_{A,B}}{O_{A,A} + O_{B,B} - O_{A,B}} \quad (1)$$

where $O_{A,A}$ is the molecule A volume, $O_{B,B}$ is the molecule B volume and $O_{A,B}$ is the overlap volume between A and B [33,34]. Color score was calculated with a color force field that included parameters for hydrogen bond donors, hydrogen bond acceptors, hydrophobic groups, anions, cations and rings [34,35].

2.4. Molecular alignment

All compounds were aligned using a benzimidazole nucleus as a template, which is a common substructure in all molecules and the minimum scaffold required for active molecules. The molecular target and action mechanism for this class of compounds are still unknown. In this paper, it is assumed that all the molecules have the same action mechanism and the same binding mode.

2.5. Tautomeric form selection

Representative tautomeric forms were selected employing the methodology called “the best predicted tautomer” [23]. This methodology is based on 3D-QSAR activity predictions and their relations with experimental activity. A first CoMFA model, including molecules with no tautomeric forms, was generated to predict the activity of molecules with tautomers. The best consensus between the experimental activity and the predicted activity for each molecule was used to select the putative active tautomeric form.

2.6. CoMFA methodology

CoMFA studies were performed with Sybyl 7.3. Steric, and electrostatic fields were obtained with Lennard-Jones potential and Coulomb potential, respectively. 3D cubic lattice with grid spacing of 2 Å and extending to 4 Å beyond the aligned molecules in all directions was used. Steric and electrostatic interactions were calculated utilizing a sp³ carbon probe atom with a charge of +1 with a distance-dependent dielectric at each lattice point. Energy truncation values of 30 kcal/mol were set for the steric and electrostatic interactions.

2.7. CoMSIA methodology

Similarity indices based on five different properties were used. Steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor descriptors were generated using the same lattice box defined in CoMFA studies, and employing the sp³ atom with charge +1 and radius of 1 Å. Similarity indices ($A_{F,k}$) for

the molecule j and the atom i at grid point q were calculated with equation (2),

$$A_{F,k}^q(j) = \sum_{i=1}^n \omega_{probe,k} \omega_{ik} e^{-\alpha r_{iq}^2} \quad (2)$$

where ω_{ik} is the actual value of the physicochemical propriety k of atom i ; $\omega_{probe,k}$ indicates the probe atom with charge +1, radius 1 Å, hydrophobicity +1, hydrogen bond donor and acceptor property +1; α is the attenuation factor; r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the test molecule. In this paper, steric indices are related to the third power of the atomic radii, electrostatic descriptors are derived from partial atomic charges, hydrophobic descriptors are derived from atom-based parameters [36], and H-bond donor and acceptor indices are obtained by a rule-based method based on experimental results [19]. The default value of 0.3 was used as the attenuation factor for the Gaussian-type distance r_{iq} .

2.8. Partial least squares analysis

Partial least squares analysis (PLS) implemented in Sybyl 7.3 was used to generate a linear regression that correlates descriptors with biological activities in pIC_{50} , cross validation being used to obtain the optimum number of the principal components. Cross validation was carried out by the Leave-One-Out (LOO) methodology, and the correlation coefficient q^2 was calculated with equation (3),

$$q^2 = 1 - \frac{\sum (Y_{pred} - Y_{exp})^2}{\sum (Y_{exp} - Y_{mean})^2} \quad (3)$$

where Y_{pred} , Y_{exp} and Y_{mean} are the values for the predicted activity, experimental activity and mean activity, respectively.

2.9. Validation

Validation was carried out considering the q^2 coefficient and predicting the activity of an external test set with 13 compounds. The trichomonocidal activities of the test set was predicted using all the CoMFA and CoMSIA models and were analyzed by regression analysis, applying the validation criteria proposed by Golbraikh and Tropsha [24,25]:

- I. $q^2 > 0.5$
- II. $R^2 > 0.6$
- III. $[(R^2 - R_0^2)/R^2] < 0.1$ or $[(R^2 - R_0^2)/R^2] < 0.1$
- IV. $0.85 \leq k \leq 1.15$ or $0.85 \leq k' \leq 1.15$

where, q^2 is the cross-validated correlation coefficient from LOO; R^2 is the correlation coefficient for experimental (y) vs. predicted (\hat{y}) activities for test set molecules; R_0^2 and $R_0'^2$ are the correlation coefficients for the regression through origin for y vs. \hat{y} and \hat{y} vs. y respectively; k and k' are the slopes for regression through origin $y^r = k\hat{y}$ and $\hat{y}^r = k'y$ and were calculated by equations (4) and (5).

$$k = \frac{\sum y_i \hat{y}_i}{\sum \hat{y}_i^2} \quad (4)$$

$$k' = \frac{\sum y_i \hat{y}_i}{\sum y_i^2} \quad (5)$$

The models that satisfied the four conditions above (I–IV) were considered highly predictive and close to being ideal models.

3. Results and discussion

A crucial step for the development of successful models of CoMFA and CoMSIA is the selection of conformers, especially when the biological target is unknown and the molecules in the study have substituents with several rotatable bonds [13]. Models based on low-energy structures have been useful in several papers [37–39]. In this paper, the minimum energy conformation for each compound obtained from conformational analysis was selected. In an alternate approach, there were employed 3D similarity methods based on the fact that ligand-receptor binding is generally characterized by a high molecular complementarity [40].

It is very important to consider tautomeric equilibrium because different tautomers have different hydrophobicity, pKa, 3D shape and electrostatic properties; furthermore, in some cases the most stable tautomer in solution is not necessarily the most biologically active, especially when the conversion barriers are low [41–43]. Most of the compounds in this paper can exist in two or more tautomeric forms. In order to select a representative tautomer, a CoMFA model was developed using benzimidazole derivatives with no tautomers. The model was used to predict the activity of the putative tautomeric forms of benzimidazole derivatives. All tautomers with the best prediction were selected [23]. The most important tautomeric forms considered are shown in Fig. 1.

3.1. CoMFA models constructed for tautomer selection approach

First, a CoMFA model 1 with 28 molecules and no tautomeric forms was constructed. These molecules were in their lowest

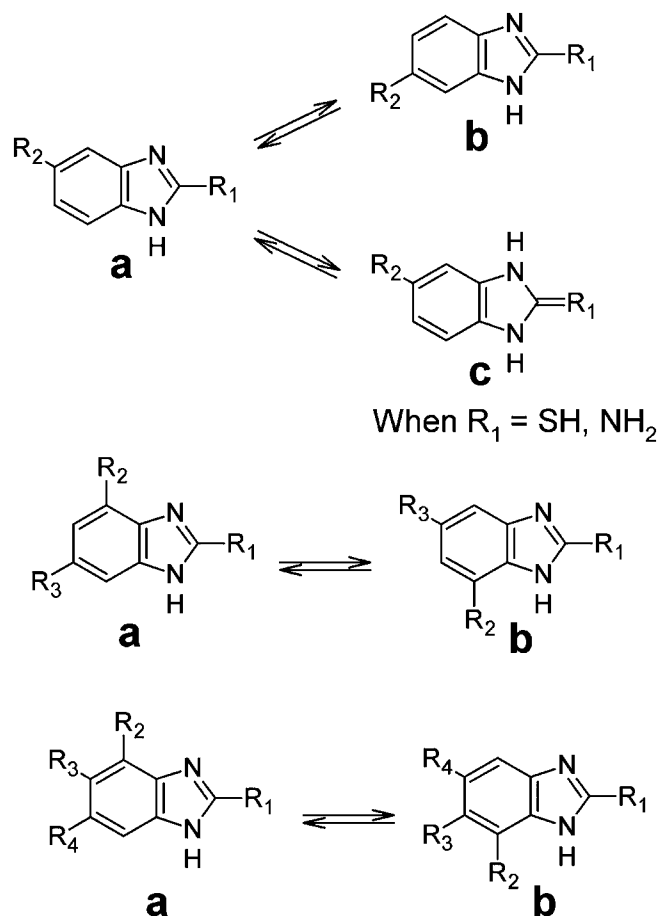


Fig. 1. Tautomeric forms considered in this paper, their identification being based on substitution patterns.

energy conformation with electrostatic potential atomic charges. The results for this model showed poor correlation and predictive ability ($r^2 = 0.443$ and $q^2 = 0.043$; Table 2). Compounds **28** and **31** were clearly responsible for the poor correlation. Then, model 2 with 26 compounds (excluding **28** and **31**) was constructed, the results showed a notable enhancement in correlation and LOO coefficient $r^2 = 0.975$ and $q^2 = 0.503$. This model was considered acceptable to predict activity values based on its q^2 value > 0.5 . Interestingly, the differences in the correlations for models based on 28 and 26 compounds are very high. The presence of compounds with behavior similar to that of **28** and **31** has been extensively studied and has been associated with different action mechanism, binding mode and flexible binding sites [44–46], and the presence of activity cliffs (molecules with small changes in structure with large differences in activity) [47,48]. In all cases, the activity of these molecules cannot be explained by conventional QSAR models. Although in this work there is no experimental evidence to explain the “outlier” behavior of compounds **28** and **31**, the presence of a bromine atom at 4- to 7-positions of the benzimidazole nucleus can be an important feature to explain the discontinuous SAR [49].

An additional feature to consider is that compounds with larger substituents at the 5- or 6-positions (i.e. **12–13**) have two low-energy rotamers with the same energy. These rotamers are oriented at the opposite sides of the plane defined by the benzimidazole nucleus. For this reason, in addition to model 2, models 3–5 clustering the molecules by different rotamers were generated (Fig. 2). The best results were observed for model 2 ($q^2 > 0.5$), which was used to estimate the activity of molecules with tautomeric forms and select the best predicted tautomer as putatively responsible for trichomonocidal activity. The values for experimental and predicted activities are shown in Table 3.

3.2. CoMFA and CoMSIA models based on minimum energy conformers

To generate CoMFA and CoMSIA models, which are based on minimum energy conformers, only compounds with no tautomers and the best predicted tautomeric forms were employed to give a total of 55 molecules. CoMFA model 6 was generated with electrostatic potential atomic charges calculated with PM3. The model had a high correlation and LOO values with $r^2 = 0.936$ and $q^2 = 0.634$, respectively (Table 4). An important fraction value of 0.475 of electrostatic contribution was observed for this model. Based on this

Table 2
Summary of CoMFA models 1–5.

Statistics	Model				
	1	2	3	4	5
q^{2a}	0.043	0.503	0.465	0.472	0.433
PRESS ^b	0.895	0.615	0.638	0.618	0.640
Compounds ^c	28	26	26	26	26
ONC ^d	1	6	6	5	5
r^{2e}	0.443	0.975	0.975	0.964	0.971
S ^f	0.683	0.138	0.137	0.161	0.144
F ^g	20.670	123.045	124.608	107.995	135.622
Steric ^h	0.518	0.501	0.509	0.520	0.555
Electrostatic ⁱ	0.482	0.499	0.491	0.480	0.445

^a Cross-validated correlation coefficient from LOO.

^b Standard error of predictions derived from LOO method.

^c Number of compounds in the training set.

^d Optimum number of principal components.

^e Non cross-validated r^2 .

^f Standard error estimate.

^g F-test value.

^h Steric field contribution.

ⁱ Electrostatic field contribution.

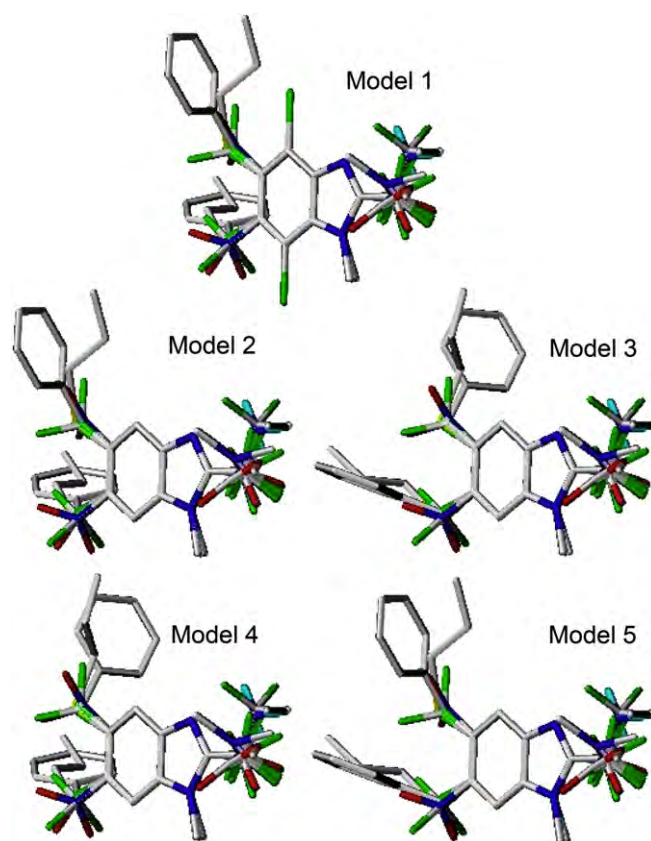


Fig. 2. Orientation of substituents at 5- to 6-positions of benzimidazole in CoMFA models 1–5. Non-polar hydrogens were removed for clarity.

observation it was decided to explore the Mulliken and Gasteiger-Hückel partial atomic charges to generate models 7 and 8, respectively. These models had high correlation ($r^2 = 0.927$ and 0.928 , respectively) and good LOO values ($q^2 = 0.602$ and 0.627 , respectively); however, model 6 gave the best values (Table 4).

The same molecules used in CoMFA model 6 were used to generate four additional models (9–12) using the CoMSIA methodology (Table 4). Thus, model 9 was constructed using steric and electrostatic fields, model 10 included donor and acceptor of hydrogen bonds fields and model 11 included a hydrophobic field. CoMSIA models 9–11 showed acceptable values of r^2 and q^2 ; particularly, model 9 had the highest values ($r^2 = 0.915$ and $q^2 = 0.617$). Model 12 was constructed using all fields, and the value of q^2 obtained increased compared with previous models. The difference in r^2 and q^2 values from models 9 and 12 was low; however the contour maps of model 12 were more informative.

3.3. CoMFA and CoMSIA studies based on conformers selected using 3D similarity

CoMFA and CoMSIA models based on minimum energy conformer selection are restricted and require that all minimum energy conformers acquire similar arrangement in 3D-space. However, a restricted space can be defined based on the fact that specific binding of small molecules to target proteins is usually similar and characterized by specific interactions and shape [40]. In this paper, ROCS methodology was used to perceive a similarity between molecules based on their three dimensional shape [34]. This methodology was used for the first time to select the best shape conformation based on query molecules for CoMFA and CoMSIA studies.

Table 3
Experimental and predicted pIC₅₀ values by CoMFA model 2.

	Experimental	Predicted	Residual		Experimental	Predicted	Residual
2a	6.35	5.72	0.63	33a^a	6.83	7.07	-0.24
2b^a		5.94	0.41	33b		6.56	0.27
3a^a	5.50	5.63	-0.13	34a^a	7.14	7.21	-0.07
3b		5.82	-0.32	34b		6.66	0.48
4a	6.63	5.60	1.03	35a	6.51	7.20	-0.69
4b^a		5.61	1.02	35b^a		6.70	-0.19
5a^a	5.64	5.61	0.03	36a^a	7.39	7.26	0.13
5b		5.92	-0.28	36b		6.83	0.56
8a	6.46	7.13	-0.67	37a^a	7.53	7.21	0.32
8b^a		6.03	0.43	37b		6.92	0.61
11a^a	4.55	4.88	-0.33	38a^a	7.07	7.19	-0.12
11b		5.35	-0.80	38b		6.92	0.15
16a^a	6.54	6.90	-0.36	39a^a	7.99	7.22	0.77
16c		5.69	0.85	39b		6.65	1.34
17a^a	6.71	6.37	0.34	40a^a	7.14	7.32	-0.18
17c		5.01	1.70	40b		6.74	0.40
18a^a	6.69	6.69	0.00	41a^a	7.69	7.30	0.39
18b		6.93	-0.24	41b		6.77	0.92
19a^a	6.79	6.68	0.11	42a^a	7.29	7.29	0.00
19b		6.92	-0.13	42b		6.75	0.54
20a^a	6.71	7.06	-0.35	43a^a	7.23	7.22	0.01
20b		7.30	-0.59	43b		6.78	0.45
20c		6.02	0.69	44a^a	7.96	7.45	0.51
21a	6.87	6.56	0.31	44b		6.89	1.07
21b^a		6.79	0.08	45a^a	8.48	7.45	1.03
21c		5.20	1.67	45b		6.96	1.52
22a	7.03	6.90	0.13	66a^a	6.67	6.65	0.02
22b^a		7.15	-0.12	66b		7.06	-0.39
23a^a	6.98	7.44	-0.46	67a^a	6.79	6.92	-0.13
23c		6.26	0.72	67b		6.98	-0.19
24a^a	5.80	6.18	-0.38	68a	7.54	6.63	0.91
24b		6.32	-0.52	68b^a		6.93	0.61
26a^a	6.72	6.56	0.16	69a^a	7.26	6.71	0.55
26b		6.00	0.72	69b		6.59	0.67
27a^a	6.57	6.44	0.13	70a^a	6.05	6.28	-0.23
27b		6.02	0.55	70b		6.47	-0.42
32a	6.16	7.06	-0.90				
32b^a		6.48	-0.32				

^a Selected tautomer based on its predicted and experimental activity.

Based on their substitution patterns, three molecules in their minimum energy conformation were used as queries in this work. The first one, **45**, was selected as query because it has the largest substituent at the 2-position and, besides it, is one of the most active compounds. The other two, **9** and **10**, were selected considering that the long chain was fixed at the 5- and 6-positions, respectively. For the other compounds with rotatable bonds shown in Table 1, the conformers were selected by considering the highest similarity value with query molecules with the same substitution pattern. It was observed that all molecules, overlaid with ROCS, showed an almost perfect alignment of the benzimidazole nucleus. Hence, this alignment was considered for the QSAR calculations.

CoMFA model 13 was constructed with conformers selected by using 3D similarity, electrostatic charges and tautomers defined by model 2. High correlation $r^2 = 0.934$ and LOO coefficient value $q^2 = 0.601$ were obtained (Table 5). These results are comparable with the minimum energy based CoMFA models. In addition, CoMSIA models 14–17 were constructed using the ROCS based conformer selection. 3D similarity based CoMSIA models were constructed using the same fields as minimum energy based models. The results showed that model 14 (with steric and electrostatic fields) and model 17 (with five fields) were the best models with the highest r^2 and q^2 values.

ROCS based models had r^2 and q^2 values very close to those of minimum energy based models. These results suggest that CoMFA models using conformer selection obtained by 3D overlay can be

useful when the QSAR models based on minimum energy conformers fail to predict molecules with several rotatable bonds.

3.4. External validation of CoMFA and CoMSIA models

Validation is a crucial step in structure–activity relationship studies. Normally, cross-validation is used to assess internal predictive power of models. However, models with high values of LOO r^2 ($q^2 > 0.5$), are not always predictive [24,25]. For this reason, it is necessary to have an external set with at least five compounds with similar kind of structures and activity values similar to training set in order to have a best validation [18,24,25]. In this paper, 13 compounds were selected as a test set and their activities were calculated using models 6–17. Furthermore, the regression plots of experimental versus predicted activities and predicted versus experimental activities were constructed to apply the conditions described in the section 2.9. The first and second conditions mean that the LOO coefficient q^2 and test set correlation coefficient R^2 should be at least 0.5 and 0.6, respectively. These conditions are necessary but not sufficient to have a highly predictive model. The third condition ($(|(R^2 - R_0^2)/R^2|) < 0.1$ or $(|(R^2 - R_0'^2)/R^2|) < 0.1$) means that at least one of the correlation coefficients for regressions through the origin R_0^2 or $R_0'^2$ should be close to R^2 ; whereas the fourth condition ($0.85 \leq k \leq 1.15$ or $0.85 \leq k' \leq 1.15$) means that at least one slope of the regression lines through the origin should be close to one. The application of these conditions suggests that our models are close to being an “ideal” model [24]. The necessary values

Table 4
Summary of statistical data and validation for CoMFA and CoMSIA models 6–12.

Statistics	CoMFA model/Charges			CoMSIA model/Charges			
	6	7	8	9	10	11	12
	Electrostatic potential	Mulliken	Gasteiger-Hückel	Electrostatic potential			
q^2	0.634	0.602	0.627	0.617	0.507	0.503	0.642
PRESS	0.468	0.488	0.472	0.474	0.527	0.535	0.454
ONC	5	5	5	4	2	3	3
r^2	0.936	0.927	0.928	0.915	0.573	0.706	0.858
S	0.196	0.209	0.208	0.223	0.491	0.412	0.286
F	142.893	124.974	125.819	134.483	34.844	40.730	102.338
Steric	0.525	0.554	0.556	0.231			0.084
Electrostatic	0.475	0.446	0.444	0.769			0.285
Donor ^a					0.596		0.227
Acceptor ^a					0.404		0.177
Hydrophobic ^b						1	0.226
R^{2c}	0.890	0.890	0.842	0.873	0.562	0.759	0.809
R_o^{2d}	0.828	0.843	0.809	0.777	0.509	0.663	0.749
k^e	1.002	0.996	1.004	0.993	0.994	0.999	0.998
$(R^2 - R_o^2)/R^2$	0.069	0.052	0.039	0.110	0.094	0.126	0.074
R'_o^{2d}	0.645	0.706	0.650	0.434	-0.790	-0.05	0.416
k'^e	0.996	1.001	0.993	1.003	0.999	0.996	0.998
$(R^2 - R'_o^2)/R'^2$	0.275	0.206	0.228	0.503	2.406	1.066	0.486

q^2 , PRESS, ONC, r^2 , S, F, steric and electrostatic have the same meaning as in Table 2.

^a Donor and acceptor of hydrogen bond fields, respectively.

^b Hydrophobic field contribution.

^c Correlation coefficient derived from predictions of test set molecules.

^d Correlation coefficients for the regression through origin for experimental vs. predicted and predicted vs. experimental activity, respectively.

^e Slopes for regression through origin from experimental vs. predicted and predicted vs. experimental, respectively.

obtained from correlation charts and the results of the application of the conditions are shown in Tables 4 and 5. An example of the regressions for test set predictions for model 6 is shown in Fig. 3. Although both plots may appear redundant, these plots could be characterized by different statistics.

CoMFA models 6–8 satisfied the validation conditions; model 6 in particular showed the best q^2 and R^2 values. In this case model 6 is in agreement that the best q^2 model has higher R^2 . On the other hand, CoMSIA models 9–12 showed acceptable q^2 values. Especially model 12, which had the highest q^2 value and the second best R^2 value; however, models 9 and 11 failed to satisfy the third condition and model 10 failed to satisfy the second condition. Therefore, only model 12 is highly predictive. For 3D similarity based models 13 through 17, all of the validation rules were satisfied with the exception of model 15. Although, models 13 through

17 have lower values of q^2 and R^2 than models 6 to 12, they better fulfill the third condition. The predicted values for the test set using the best models are shown in the Table 6.

3.5. CoMFA and CoMSIA contour maps

Contour maps for the best CoMFA and CoMSIA models (6 and 12, respectively) are shown in Fig. 4. The discussion of the contour maps is presented in the following six sections.

3.5.1. Steric CoMFA contour maps

Steric contour maps (Fig. 4a) are shown in green when the activity increases with large or bulky substituents. The opposite is true for yellow contours. There is a yellow contour close to the 1-position in the benzimidazole nucleus, which suggests that trichomonocidal activity decreases with methyl substitution. This trend is observed for all benzimidazole derivatives substituted at other positions by small groups (1, 2, 4–7, 14–21, 23–26, and 29, 30), where the 1-methylbenzimidazole derivatives 6 and 7 are the least active compounds ($pIC_{50} = 5.39$ and 5.27, respectively). Indeed, the most active compounds ($pIC_{50} > 7$) have hydrogen at the 1-position with the exception of compounds 49, 56, 57, and 61, where the substituent at the 2-position plays an important role in activity. This trend is different for compounds 8–10, 12, and 13, where the larger substituents at the 5- and 6-positions determine the differences in activity. A favorable steric contour is found at the 2-position of the benzimidazole nucleus. This region is associated with the high activity of the 2-alkylthiobenzimidazole derivatives 32–35, 37–40, 42–45, and the 2-ethyl ester derivatives 49, 57, and 61 (pIC_{50} values 6.16–8.48). This group contains most of the active compounds in the data set. A yellow steric unfavorable region at the 5-position and a green contour close to the 6-position indicate that compounds substituted for at the 6-position are more active than those substituted for at the 5-position. This is observed with 1-methylbenzimidazole derivatives 12, 13, 47, 51, 63, and 64. However, this trend is not so clear with compounds 6 and 7 because they have almost the same activity. The opposite trend, on the other

Table 5
Summary of statistical data and validation for CoMFA and CoMSIA models 13–17.

Statistics	CoMFA model		CoMSIA model		
	13	14	15	16	17
q^2	0.601	0.565	0.482	0.506	0.585
PRESS	0.489	0.510	0.551	0.538	0.484
ONC	5	5	4	4	2
r^2	0.934	0.914	0.669	0.748	0.805
S	0.198	0.227	0.441	0.385	0.332
F	139.397	104.593	25.296	37.061	107.175
Steric	0.499	0.245			0.095
Electrostatic	0.501	0.755			0.307
Donor			0.603		0.247
Acceptor			0.397		0.121
Hydrophobic				1	0.230
R^2	0.825	0.729	0.689	0.678	0.760
R_o^2	0.795	0.692	0.589	0.644	0.726
k	0.995	0.993	0.996	1.009	0.995
$(R^2 - R_o^2)/R^2$	0.036	0.051	0.145	0.050	0.045
R'_o^2	0.625	0.301	-0.54	0.135	0.426
k'	1.002	1.002	0.998	0.986	1.001
$(R^2 - R'_o^2)/R'^2$	0.242	0.587	1.784	0.801	0.439

All values have the same meaning as they do in Tables 3, 4.

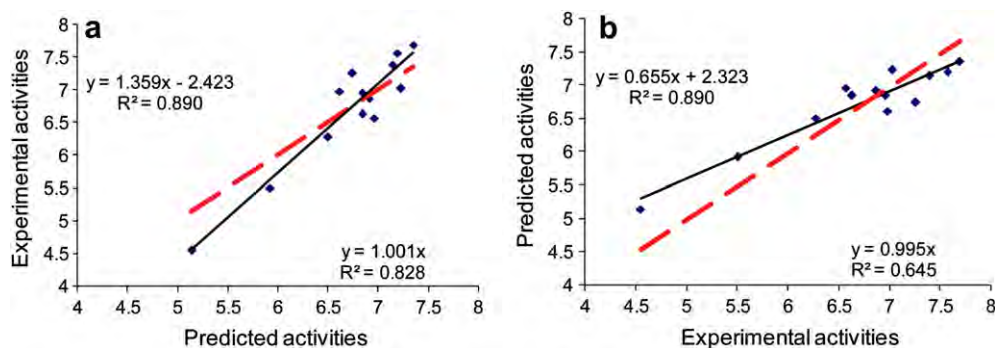


Fig. 3. Validation for CoMFA model 6 using the test set. (a) Experimental vs. predicted activities. (b) Predicted vs. experimental activities.

hand, is observed for compounds **9** and **10**. The difference in activity between the latter two compounds (pIC_{50} values of 6.7 and 5.59, respectively) can explain the yellow contour at the border of the 6-position, thus suggesting that large substituents at this position decreases the activity.

3.5.2. Electrostatic CoMFA contour maps

Electrostatic contour maps are shown in red when the activity increases with negative atomic charges and in blue when the activity decreases (Fig. 4b). In general, red contour maps are close to heteroatoms such as nitrogen and oxygen, whose partial atomic charges are highly negative. Three negative favorable red regions are observed close to the substituents at the 2-position. One of them is close to the oxygen in the carbonyl group of compounds **47**, **49–52**, **55–58**, and **60–61**. These compounds showed moderate activity for amide derivatives (pIC_{50} 6.37–7.12) and strong activity for ester derivatives (pIC_{50} 7.07–7.72). A second one is extended from the amide nitrogen and the alkyl oxygen, for the same group of compounds (**47**, **49–52**, **55–58**, and **60–61**), to the nitrogen of the amino group in the high active compounds **34**, **37–40**, and **42–43** ($pIC_{50} > 7$). The last red contour is at the extreme of the 2-position. It can be associated with the nitro group in two of the more active compounds **44** and **45** ($pIC_{50} = 7.96$ and 8.48 , respectively). Besides, a blue contour at the 2-position is shown to be close to the aliphatic region of the moderately to highly active compounds **32–35**, **37–40**, and **42–45** (pIC_{50} values 6.16–8.48). The same contour is also close to the ethyl chain in the potent ethyl ester derivatives **49**, **57**, and **61** ($pIC_{50} > 7$). Three red contours close to the 5- and 6-positions suggest that negative charged groups such as halogen atoms increase the activity as compared to the hydrogen atom. Furthermore, larger substituents are associated with a blue contour, which suggests the favorable effect of the aliphatic portion in the substituent propylthio **8–10** ($pIC_{50} = 5.59–6.70$), and alkyl amides **68** and **70** ($pIC_{50} = 6.05–7.54$) as compared with the electron rich benzoyl substituent in the less active compounds **12**, **13** ($pIC_{50} = 4.53$ and 4.97 , respectively).

3.5.3. Steric CoMSIA contour maps

Coloring in CoMSIA steric contour maps (Fig. 4c) is similar to the color schemes for the CoMFA maps; however, although both maps have similar appearance, that of CoMSIA offer a greater level of detail. An unfavorable yellow CoMSIA contour is observed close to the 1-position of the benzimidazole with a better definition than that of CoMFA and suggests the favorable effect of hydrogen substitution. A green contour at the 2-position is shown, suggesting the favorable substitution effect in this position with large substituents. In agreement with CoMFA, a prominent unfavorable yellow contour is observed at the 5- and 6-positions, indicating the unfavorable effect of the substituents at the 5-position and the large substituents at the 6-position.

3.5.4. Electrostatic CoMSIA contour maps

The colors of the electrostatic contour maps in the CoMSIA models (Fig. 4d) have the same meaning as those of the CoMFA models. Similar to CoMFA, a blue contour map is observed at the 2-position of the benzimidazole. However, unlike CoMFA, the CoMSIA contour is larger and covers the aliphatic region of **32–35**, **37–40** and **42–43**, and the deactivated aromatic ring of **44** and **45**. Also similar to CoMFA, a second blue contour is shown at the 5- and 6-positions. This contour is close to the aromatic ring of inactive compound **13** ($pIC_{50} = 4.97$) and to the oxygen in carbonyl groups of **12**, **66**, **67**, and **70** ($pIC_{50} = 4.53–6.79$, median $pIC_{50} = 6.36$).

3.5.5. Hydrogen bond donor and acceptor CoMSIA contour maps

Areas favored by donors and acceptors are shown in cyan and magenta respectively; unfavorable areas are in purple and red, respectively (Fig. 4e). At the 1-position of the benzimidazole nucleus the contour is indicated in cyan, which suggests favorable contributions of hydrogen in the 1*H*-benzimidazole derivatives. This observation is in agreement with the yellow CoMFA steric contour generated by the 1-methylbenzimidazole derivatives. This contour can be associated with the higher activity of most of the 1*H*-benzimidazole derivatives in the database. A purple unfavorable hydrogen bond donor contour map is close to the substituents at the 2-position. The purple contour is associated with the decrement in the activity of the hydroxyl substituents of **32** and **33** ($pIC_{50} < 7$). Furthermore, this contour is close to the *N*-H of the amide type of compounds that are less active than are the ester

Table 6

Training set predictions using models 6, 12, 13 and 17.

	Model								
		6		12		13		17	
	Exp. ^a	Pred. ^b	Res. ^c	Pred.	Res.	Pred.	Res.	Pred.	Res.
3a	5.50	5.92	−0.42	6.18	−0.68	5.88	−0.38	6.09	−0.59
11a	4.55	5.14	−0.59	5.4	−0.85	5.17	−0.62	5.33	−0.78
22b	7.03	7.23	−0.2	7.32	−0.29	7.24	−0.21	7.16	−0.13
27a	6.57	6.96	−0.39	6.71	−0.14	6.89	−0.32	6.56	0.01
36a	7.39	7.14	0.25	7.36	0.03	7.41	−0.02	7.51	−0.12
41a	7.69	7.35	0.34	7.52	0.17	7.28	0.41	7.42	0.27
46	6.96	6.84	0.12	6.81	0.15	7.04	−0.08	6.91	0.05
48	6.63	6.84	−0.21	6.97	−0.34	6.8	−0.17	7.25	−0.62
53	7.57	7.19	0.38	7.09	0.48	7.26	0.31	7.13	0.44
54	6.87	6.91	−0.04	6.81	0.06	7.18	−0.31	6.92	−0.05
59	6.98	6.61	0.37	6.74	0.24	6.3	0.68	6.77	0.21
65	6.27	6.50	−0.23	6.25	0.02	6.62	−0.35	6.46	−0.19
69a	7.26	6.74	0.52	6.53	0.73	6.8	0.46	6.38	0.88

^a Experimental activity.

^b Predicted activity.

^c Residual.

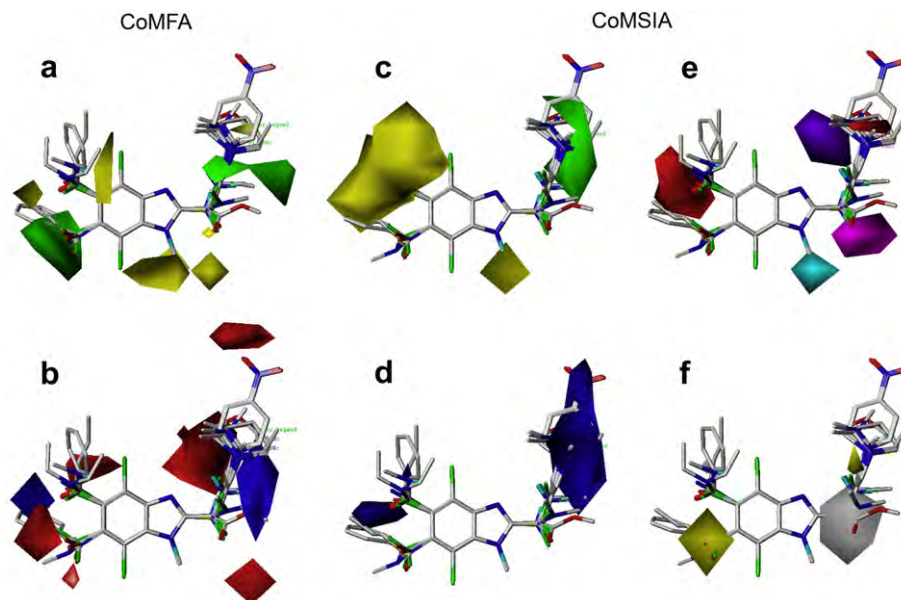


Fig. 4. CoMFA and CoMSIA STDEV*COEFF contour maps for the best models. CoMFA model 6: (a) Sterically favored areas are in green, and sterically unfavored areas are in yellow. (b) Negative charge favored areas are in red and unfavored areas are in blue. CoMSIA Model 12: The colors in (c) and (d) have the same meaning as do CoMFA contour maps. (e) Donor and acceptor favored areas are in cyan and magenta, respectively, and donor and acceptor unfavored areas are in purple and red, respectively. (f) Hydrophobic favored areas are in yellow and unfavored areas in gray. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

analogues. A magenta favorable hydrogen bond acceptor contour is shown at the 2-position of benzimidazole and is associated with the carbonyl group in the amide and ester substituents. This contour is in agreement with the CoMFA red electrostatic contour in the same area (Fig. 4b). A hydrogen bond acceptor unfavorable contour in red is shown at the 2-position of the benzimidazole. This contour can be associated with an increment in activity when the linker carbon chain, between acceptor groups and the benzimidazole nucleus, changes from two to three methylene groups in compounds **34**, **35** and **38**, **39** (pIC_{50} 6.51–7.14 to 7.07–7.99, respectively). A second red contour is close to the carbonyl group at the 5-position in some amides and benzoyl derivatives.

3.5.6. Hydrophobic CoMSIA contour maps

Hydrophobic contour maps (Fig. 4f) are shown in yellow when the activity increases with hydrophobic groups and in gray when the activity decreases. A gray contour map is shown at the 2-position of the benzimidazole nucleus generated by the poor activity of trifluoromethyl and pentafluoroethyl substituted benzimidazoles (pIC_{50} 4.53–6.98, median pIC_{50} = 6.24). A small favorable yellow contour map is observed at the 2-position, close to the aliphatic portion of the substituents in **38**, **39**. A second favorable yellow hydrophobic contour map is close to the 6-position of the benzimidazole nucleus. The contour can be explained by the presence of the halogen atom, which in most cases produces more active compounds. This contour is in agreement with the green contour at the same position in CoMFA (Fig. 4a).

4. Conclusions

The benzimidazole derivatives included in this study had a reasonable SAR and strong quantitative correlations. However, some exceptions were observed in compounds **28** and **31**, which have small changes in structure but large changes in activity.

The CoMFA and CoMSIA models were generated and showed good r^2 and q^2 values in almost all models. Eight models revealed a good response to test set validation. CoMFA models based on minimum energy conformations presented better test set

predictions than did CoMSIA. However, the CoMSIA models offered valuable information, since they explored additional kinds of fields and in some cases gave clearer contour maps.

The CoMFA and CoMSIA models, generated using 3D similarity conformer selection, showed good results and complemented models based on minimum energy. These models could be useful to estimate the activity of some molecules with high rotational degrees of freedom on substituents at the 2-, 5-, and 6-positions.

The best CoMFA model (model 6, r^2 = 0.936, q^2 = 0.634), based on the steric and electrostatic fields, included electrostatic potential atomic charges and minimum energy conformers. The best CoMSIA model (model 12, r^2 = 0.858, q^2 = 0.642), based on five fields, included the same charge type and energy conformations as the CoMFA model 6. The analysis of the contour maps for the best CoMFA and CoMSIA models indicates that hydrogen at the 1-position seems to have an important role in activity. Large groups at the 2-position with hydrophobic segments generate potent compounds. Also, heteroatoms at the 2-position can generate regions with negative charges that could act as hydrogen bond acceptors involving the binding site. In general, substituents at the 6-position are more favored than are substituents at the 5-position. Substitution at the 6-position with chlorine or bromine atom is favored over other substitutions in the current data set.

These models were used in our group to design and screen a new series of benzimidazole derivatives that are currently being synthesized, the results of which will be reported in further publications.

Acknowledgements

This work was supported by project 80093 from CONACyT. J.P.-V. acknowledges CONACyT scholarship 173896/173896 and Carso Health Institute for the PhD student fellowship S0710443. J.L.M.-F. acknowledges the State of Florida for funding. The authors are grateful to OpenEye Scientific Software, Inc., for providing ROCS.

References

- [1] Global Prevalence and Incidence of Selected Curable Sexually Transmitted Infections: Overview and Estimates. World Health Organization, Geneva, 2001.
- [2] J.R. Schwebke, Update of trichomoniasis, *Sex. Transm. Infect.* 78 (2002) 378–379.
- [3] P. Upcroft, J.A. Upcroft, Drug targets and mechanisms of resistance in the anaerobic protozoa, *Clin. Microbiol. Rev.* 14 (2001) 150–164.
- [4] S.L. Cudmore, K.L. Delgaty, S.F. Hayward-McClelland, D.P. Petrin, G.E. Garber, Treatment of infections caused by metronidazole-resistant *Trichomonas vaginalis*, *Clin. Microbiol. Rev.* 17 (2004) 783–793.
- [5] M. Boiani, M. Gonzalez, Imidazole and benzimidazole derivatives as chemotherapeutic agents, *Mini Rev. Med. Chem.* 5 (2005) 409–424.
- [6] G. Navarrete-Vázquez, R. Cedillo, A. Hernández-Campos, L. Yépez, F. Hernández-Luis, J. Valdez, R. Morales, R. Cortés, M. Hernández, R. Castillo, Synthesis and antiparasitic activity of 2-(trifluoromethyl)-benzimidazole derivatives, *Bioorg. Med. Chem. Lett.* 11 (2001) 187–190.
- [7] J. Valdez, R. Cedillo, A. Hernández-Campos, L. Yépez, F. Hernández-Luis, G. Navarrete-Vázquez, A. Tapia, R. Cortés, M. Hernández, R. Castillo, Synthesis and antiparasitic activity of 1*H*-benzimidazole derivatives, *Bioorg. Med. Chem. Lett.* 12 (2002) 2221–2224.
- [8] G. Navarrete-Vázquez, L. Yépez, A. Hernández-Campos, A. Tapia, F. Hernández-Luis, R. Cedillo, J. González, A. Martínez-Fernández, M. Martínez-Gruero, R. Castillo, Synthesis and antiparasitic activity of albendazole and mebendazole analogues, *Bioorg. Med. Chem.* 11 (2003) 4615–4622.
- [9] G. Navarrete-Vázquez, M.D. Rojano-Vilchis, L. Yépez-Mulia, V. Meléndez, L. Gerena, A. Hernández-Campos, R. Castillo, F. Hernández-Luis, Synthesis and antiprotozoal activity of some 2-(trifluoromethyl)-1*H*-benzimidazole bioisosteres, *Eur. J. Med. Chem.* 41 (2006) 135–141.
- [10] D. Valdez-Padilla, S. Rodríguez-Morales, A. Hernández-Campos, F. Hernández-Luis, L. Yépez-Mulia, A. Tapia-Contreras, R. Castillo, Synthesis and antiprotozoal activity of novel 1-methylbenzimidazole derivatives, *Bioorg. Med. Chem.* 17 (2009) 1724–1730.
- [11] M. Andrzejewska, L. Yépez-Mulia, R. Cedillo-Rivera, A. Tapia, L. Vilpo, J. Vilpo, Z. Kazimierzczuk, Synthesis, antiprotozoal and anticancer activity of substituted 2-trifluoromethyl- and 2-pentafluoroethylbenzimidazoles, *Eur. J. Med. Chem.* 37 (2002) 973–978.
- [12] M. Andrzejewska, L. Yépez-Mulia, A. Tapia, R. Cedillo-Rivera, A.E. Laudy, B.J. Starosciak, Z. Kazimierzczuk, Synthesis, and antiprotozoal and antibacterial activities of *S*-substituted 4,6-dibromo- and 4,6-dichloro-2-mercaptobenzimidazoles, *Eur. J. Pharm. Sci.* 21 (2004) 323–329.
- [13] F. Ooms, Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry, *Curr. Med. Chem.* 7 (2000) 141–158.
- [14] W.L. Jorgensen, The many roles of computation in drug discovery, *Science* 303 (2004) 1813–1818.
- [15] J.L. Medina-Franco, F. López-Vallejo, R. Castillo, Diseño de fármacos asistido por computadora, *Educ. Quím* 17 (2006) 452–457.
- [16] W.L. Jorgensen, Efficient drug Lead discovery and Optimization, *Acc. Chem. Res.* 42 (2009) 724–733.
- [17] R.D. Cramer, D.E. Patterson, J.D. Bunce, Comparative molecular field analysis. 1. Effect of shape on binding of steroids to carrier proteins, *J. Am. Chem. Soc.* 110 (1988) 5959–5967.
- [18] H. Kubinyi, Comparative molecular field analysis (CoMFA). in: P. v. R. Schleyer, N.L. Allinger, T. Clark, J. Gasteiger, P.A. Kollman, H.F. Schaefer III, P.R. Schreiner (Eds.), *The Encyclopedia of Computational Chemistry*. John Wiley & Sons, Chichester, 1998, pp. 448–460.
- [19] G. Klebe, U. Abraham, Comparative Molecular Similarity Index Analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries, *J. Comput. Aided Mol. Des* 13 (1999) 1–10.
- [20] O. Temiz-Arpaci, B. Tekiner-Gulbas, I. Yildiz, E. Aki-Sener, I. Yalcin, 3D-QSAR analysis on benzazole derivatives as eukaryotic topoisomerase II inhibitors by using comparative molecular field analysis method, *Bioorg. Med. Chem.* 13 (2005) 6354–6359.
- [21] P.D. Patel, M.R. Patel, N. Kaushik-Basu, T.T. Talele, 3D QSAR and molecular docking studies of benzimidazole derivatives as hepatitis C virus NS5B polymerase inhibitors, *J. Chem. Inf. Model.* 48 (2008) 42–55.
- [22] M.L. López-Rodríguez, M. Murcia, B. Benhamú, A. Viso, M. Campillo, L. Pardo, Benzimidazole derivatives. 3. 3D-QSAR/CoMFA model and computational simulation for the recognition of 5-HT4 receptor antagonists, *J. Med. Chem.* 45 (2002) 4806–4815.
- [23] F. López-Vallejo, J.L. Medina-Franco, A. Hernández-Campos, S. Rodríguez-Morales, L. Yépez, R. Cedillo, R. Castillo, Molecular modeling of some 1*H*-benzimidazole derivatives with biological activity against *Entamoeba histolytica*: a comparative molecular field analysis study, *Bioorg. Med. Chem.* 15 (2007) 1117–1126.
- [24] A. Golbraikh, A. Tropsha, Beware of $q(2)!$, *J. Mol. Graph. Model.* 20 (2002) 269–276.
- [25] A. Golbraikh, A. Tropsha, Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection, *J. Comput.-Aided Mol. Des* 16 (2002) 357–369.
- [26] Spartan, Version 02, Wavefunction Inc., Irvine, CA.
- [27] R.S. Mulliken, Electronic population analysis on LCAO-MO molecular wave functions, *J. Chem. Phys.* 23 (1955) 1833–1840.
- [28] S.R. Cox, D.E. Williams, Representation of the molecular electrostatic potential by a net atomic charge model, *J. Comput. Chem.* 2 (1981) 304–323.
- [29] J.J.P. Stewart, Optimization of parameters for semiempirical methods I. Method, *J. Comput. Chem.* 10 (1989) 209–220.
- [30] J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges, *Tetrahedron* 36 (1980) 3219–3228.
- [31] Sybyl, Version 7.3, Tripos Inc., St. Louis, MO.
- [32] Rocs, Version 2.3.1, OpenEye Scientific Software Inc., Santa Fe, NM.
- [33] P. Willett, J.M. Barnard, G.M. Downs, Chemical similarity searching, *J. Chem. Inform. Comput. Sci.* 38 (1998) 983–996.
- [34] T.S. Rush, J.A. Grant, L. Mosyak, A. Nicholls, A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction, *J. Med. Chem.* 48 (2005) 1489–1495.
- [35] M.J. Sykes, M.J. Soricich, J.O. Miners, Molecular modeling approaches for the prediction of the nonspecific binding of drugs to hepatic microsomes, *J. Chem. Inf. Model.* 46 (2006) 2661–2673.
- [36] V.N. Viswanadhan, A.K. Ghose, G.R. Revankar, R.K. Robins, Atomic physico-chemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics, *J. Chem. Inf. Comput. Sci.* 29 (1989) 163–172.
- [37] S.S. Kulkarni, A.H. Newman, W.J. Houlihan, Three-dimensional quantitative structure-activity relationships of Mazindol analogues at the dopamine transporter, *J. Med. Chem.* 45 (2002) 4119–4127.
- [38] V. Ravichandran, R.K. Agrawal, Predicting anti-HIV activity of PETT derivatives: CoMFA approach, *Bioorg. Med. Chem. Lett.* 17 (2007) 2197–2202.
- [39] A.K. Chakraborti, B. Gopalakrishnan, M.E. Sobhia, A. Malde, Comparative molecular field analysis (CoMFA) of phthalazine derivatives as phosphodiesterase IV inhibitors, *Bioorg. Med. Chem. Lett.* 13 (2003) 2473–2479.
- [40] G. Klebe, Recent developments in structure-based drug design, *J. Mol. Med.* 78 (2000) 269–281.
- [41] Y.C. Martin, Let's not forget tautomers, *J. Comput.-Aided Mol. Des* 23 (2009) 693–704.
- [42] A.R. Katritzky, C.D. Hall, B. El-Dien, M. El-Gendy, B. Draghici, Tautomerism in drug discovery, *J. Comput.-Aided Mol. Des* 24 (2010) 475–484.
- [43] F. Milletti, A. Vulpetti, Tautomer preference in PDB complexes and its impact on structure-based drug discovery, *J. Chem. Inf. Model.* 50 (2010) 1062–1074.
- [44] K.H. Kim, Outliers in SAR and QSAR: is unusual binding mode a possible source of outliers? *J. Comput.-Aided Mol. Des* 21 (2007) 63–86.
- [45] K.H. Kim, Outliers in SAR and QSAR: 2. Is a flexible binding site a possible source of outliers? *J. Comput.-Aided Mol. Des* 21 (2007) 421–435.
- [46] M.T. Sisay, L. Peltason, J. Bajorath, Structural interpretation of activity cliffs revealed by systematic analysis of structure-activity relationships in analog series, *J. Chem. Inf. Model.* 49 (2009) 2179–2189.
- [47] G.M. Maggiora, On outliers and activity cliffs - Why QSAR often disappoints, *J. Chem. Inf. Model.* 46 (2006) 1535.
- [48] J.L. Medina-Franco, K. Martínez-Mayorga, A. Bender, R.M. Marin, M.A. Giulianotti, C. Pinilla, R.A. Houghten, Characterization of activity Landscapes using 2D and 3D similarity methods: consensus activity cliffs, *J. Chem. Inf. Model.* 49 (2009) 477–491.
- [49] J. Pérez-Villanueva, R. Santos, A. Hernández-Campos, M.A. Giulianotti, R. Castillo, J.L. Medina-Franco, Towards a systematic characterization of the antiprotozoal activity landscape of benzimidazole derivatives, *Bioorg. Med. Chem.* 18 (2010) 7380–7391.



Antiprotozoal activity of proton-pump inhibitors

Jaime Pérez-Villanueva^a, Antonio Romo-Mancillas^a, Alicia Hernández-Campos^a, Lilián Yépez-Mulia^b, Francisco Hernández-Luis^a, Rafael Castillo^{a,*}

^aFacultad de Química, Departamento de Farmacia, UNAM, México DF 04510, Mexico

^bUnidad de Investigación Médica en Enfermedades Infecciosas y Parasitarias, IMSS, México DF 06720, Mexico

ARTICLE INFO

Article history:

Received 6 August 2011

Revised 6 October 2011

Accepted 7 October 2011

Available online 14 October 2011

Keywords:

Benzimidazole

Entamoeba histolytica

Giardia intestinalis

Proton-pump inhibitors

Trichomonas vaginalis

ABSTRACT

Parasitic diseases are still a major health problem in developing countries. In our effort to find new antiparasitic agents, in this Letter we report the in vitro antiprotozoal activity of omeprazole, lansoprazole, rabeprazole and pantoprazole against *Trichomonas vaginalis*, *Giardia intestinalis* and *Entamoeba histolytica*. Molecular modeling studies were an important tool to highlight the potential antiprotozoal activity of these drugs. Experimental evaluations revealed a strong activity for all compounds tested. Rabeprazole and pantoprazole were the most active compounds, having IC₅₀ values in the nanomolar range, which were even better than metronidazole, the drug of choice for these parasites.

© 2011 Elsevier Ltd. All rights reserved.

Parasitic infections still represent a major public health problem in developing countries. Mucosal infections by protozoa affect more than a billion people every year.¹ The most important protozoan infections include giardiasis, amoebiasis and trichomonosis, whose causal agents are *Giardia intestinalis*, *Entamoeba histolytica* and *Trichomonas vaginalis*, respectively.¹ According to the World Health Organization (WHO), 280 million giardiasis cases are estimated each year. This infection causes diarrhea and associated symptoms. Its incidence in children is very high, producing severe growth retardation and low development.^{1,2} Likewise, WHO estimated that *E. histolytica* causes 50 million amoebiasis cases and 100,000 deaths each year. The principal associated symptoms of amoebiasis are diarrhea and dysentery. When this protozoan penetrates the intestinal mucosa and migrates to other organs, it can be mortal.^{1,3} In addition to intestinal infections, the genitourinary infection caused by *T. vaginalis* is estimated to be more than 180 million new cases annually.¹ In women, trichomonosis can cause severe inflammation of the genital tract which has been associated with preterm labor, low-birth weight, sterility, cervical cancer and a predisposition to HIV infection.^{4,5} The drug of choice for the treatment of these infections is metronidazole (MTZ), which has been used for more than 40 years. Nevertheless, its side effects and the development of resistant strains have stimulated our interest to search for new antiprotozoal agents.^{1,5} Since the previous report of the giardicidal and trichomonocidal activity of benzimidazole 2-methylcarbamates (BC), albendazole (ABZ) and

mebendazole (MBZ)^{6–9}; our research group and others have been working in the synthesis and evaluation of benzimidazole derivatives as antiprotozoals.^{10–17} Interestingly, we found that benzimidazole derivatives that are not BC do not inhibit β -tubulin polymerization as antiparasitics BC do.^{10,11} In our efforts to have information about the structural requirements that a benzimidazole derivative must have for antiprotozoal activity, we did QSAR studies based on CoMFA and CoMSIA methodologies,^{18,19} and SAR studies based on the emerging concept of activity landscape with SAS maps (Structure–Activity Similarity maps).^{20,21} As a result of these computational studies it was found that substituents at positions 2, 5 and 6 are important for antiprotozoal activity.^{18–21} Another kind of benzimidazole derivatives used in therapeutics are the proton pump inhibitors (PPIs) omeprazole, lansoprazole, pantoprazole and rabeprazole (Fig. 1).²² These drugs have been extensively used for the treatment of peptic ulcer, gastroesophageal reflux and the eradication of *Helicobacter pylori* in combination with antibiotics.²³ Interestingly, PPIs shared several structural features with antiprotozoal benzimidazole derivatives.

Based in the 'Similarity principle', it is generally accepted that similar molecules have similar activities; however, this is not always true. It has been found, in some cases, that similar molecules possess different activity (*activity cliffs*).²⁴ This is not surprising since molecular recognition plays a crucial role in determining activity. The *activity cliffs* become very important in drug discovery when small changes in structure lead to favorable changes in activity. In both cases, when active compounds are known, such as antiprotozoal benzimidazole derivatives, the evaluation of similar molecules could result in valuable active compounds; especially,

* Corresponding author. Tel.: +525 56 22 52 87; fax: +525 56 22 53 29.

E-mail address: rafaelc@servidor.unam.mx (R. Castillo).

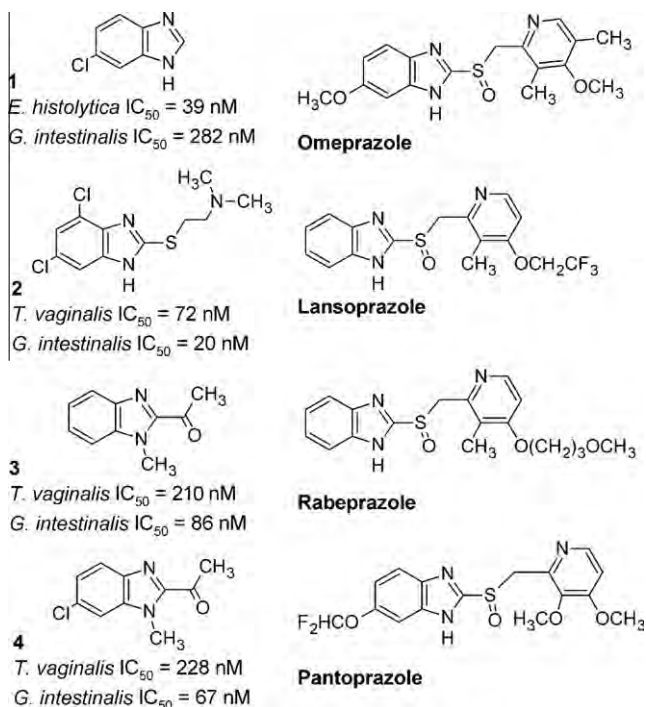


Figure 1. Some benzimidazole derivatives with antiprotozoal activity and PPIs.

when new candidates are widely used in therapeutics having known safety.²⁵ Hence, based on the structural similarities between some antiprotozoal benzimidazole derivatives and the PPIs, the antiprotozoal activity of the latter was tested against *T. vaginalis*, *G. intestinalis* and *E. histolytica*.

The benzimidazole nucleus is the common scaffold for several derivatives with strong activity against *E. histolytica*, *G. intestinalis* and *T. vaginalis*, that is, **1–4**, as well as for PPIs (Fig. 1). An additional common feature between both type of compounds is their side chain similarities at position 2, where nitrogen and sp² oxygen can play an important role as hydrogen bond acceptors. Also, polar hydrogen at position 1 and nitrogen at position 3 can be a potential hydrogen bond donor and acceptor respectively.

To visualize 3D similarities between compounds **1–4** and PPIs, lower energy structures were obtained using molecular modeling; conformational analysis and equilibrium geometries were performed with Spartan'02 software.^{26,27} The molecules were aligned using Sybyl 7.3 software, considering benzimidazole nucleus and common hydrogen bond acceptor groups at position 2.²⁸ The results showed stable conformations with similar 3D arrangements (Fig. 2). Furthermore, triangle type distances were measured to compare the reference compounds with PPIs. Reference points are located on a centroid at the benzenoid ring and at common hydrogen bond acceptors at position 2. Indeed, triangle type distances measured in compounds **2** and **3** are almost the same for PPIs.

Moreover, molecular electrostatic potential calculations were performed using the same structures. Results showed high electrostatic similarity between reference molecules and PPIs (Fig. 3). Isocontours close to nitrogen at position 3 are common in all benzimidazole derivatives. Also, isocontours at position 2 near the nitrogen and sp² oxygen are common in molecules **2–4** and PPIs. Furthermore, similar isocontours are shown at position 5(6) in omeprazole and pantoprazole as compared with compounds **1**, **2** and **4**. These results suggest that PPIs would probably show antiprotozoal activity involving the same binding site as that of compounds **1–4**, based on the fact that ligand–receptor binding

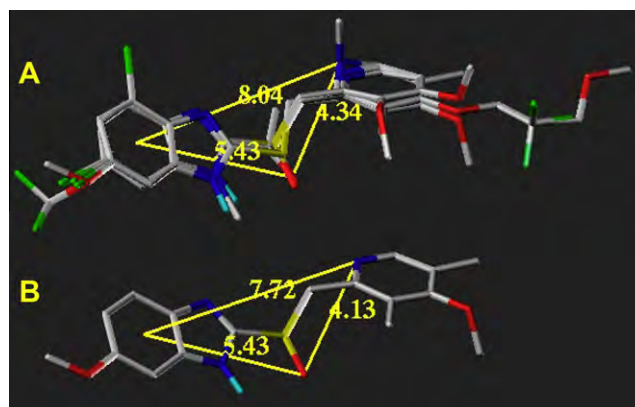


Figure 2. (A) Molecular overlap for compounds **1–4** and PPIs. Triangle type distances were measured for reference molecules **2** and **3** considering some of their putative pharmacophoric elements; (B) Triangle type distances for omeprazole.

is generally characterized by a high molecular complementarity and common interactions.²⁹

Compounds were isolated by extraction from their commercial presentation.³⁰ The chemical structures of the extracted compounds were established by ¹H Nuclear Magnetic Resonance spectroscopy and Mass spectrometry data.³¹ The in vitro antiprotozoal activity of PPIs was then tested following a method previously described^{10–17}; results are shown in Table 1. All compounds showed good activity against the three protozoa in the nanomolar range. PPIs were 1.9–3.1 times more active than metronidazole against *T. vaginalis* and 12.8–78.1 more active against *G. intestinalis*. Besides, rabeprazole and pantoprazole were 14.8 and 134.6 times more active than MTZ respectively against *E. histolytica*. It is worth noting that pantoprazole showed very good activity against the three anaerobic protozoa tested.

Although omeprazole and lansoprazole have different groups and substitution patterns, there are no significant differences in their activity against *T. vaginalis*, while against *G. intestinalis* and *E. histolytica* the differences are low. The differences in activity are more remarkable when lansoprazole and rabeprazole are compared, thus indicating that the substituent at position 2 of the benzimidazole nucleus has an important role in the activity against the three parasites. Furthermore, the increment in activity of pantoprazole against all parasites, as compared with omeprazole, suggests an important electrostatic contribution caused by the difluoromethoxy substituent at position 5(6) of the benzimidazole nucleus.

Considering the novel antiprotozoal activity for PPIs, these agents are candidates for drug repurposing, that is a new use. In favor of this is the fact that they have known pharmacokinetics and safety profiles and are often approved by regulatory agencies for human use.²⁵ Drug repurposing is an increasingly popular strategy in pharmaceutical companies because is a proven short cut between the lab and the clinic. Hence, repurposing is a good strategy to gain access for new therapies; specially, for neglected diseases like parasitic infections.^{25,32–34}

In summary, considering that parasitic diseases caused by protozoa are still an important public health problem, during the last years our efforts have been centered on finding new antiprotozoal agents. The benzimidazole derivatives synthesized and tested by our group, and by other research groups, provide important structural information to gain access to new active compounds. Some of these compounds have similar structural features as compared with PPIs. The study of these similarities aided by molecular modeling tools has led us to discover their novel antiprotozoal activity. PPIs showed very good antiprotozoal activity in the nanomolar

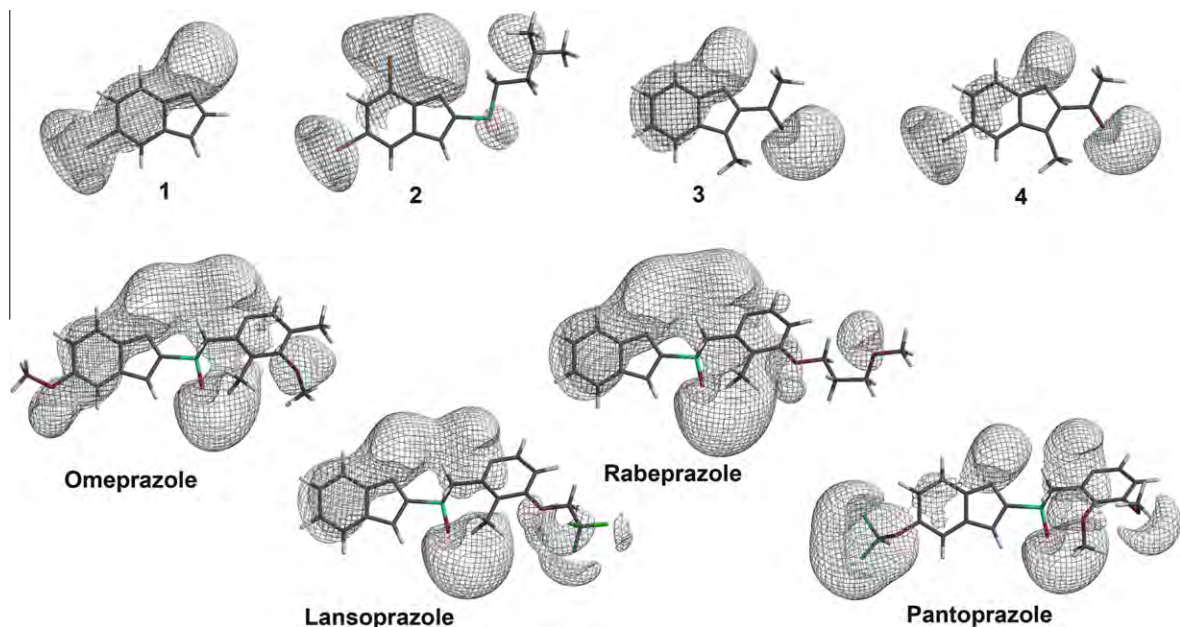


Figure 3. Molecular electrostatic potential for derivatives 1–4 and PPIs calculated at Hartree-Fock level (HF/6–31G^{*}//HF/6–31G^{*}) using an isovalue of –10 kcal/mol.

Table 1

In vitro susceptibility of *T. vaginalis*, *G. intestinalis* and *E. histolytica* to commercially available PPIs

	<i>T. vaginalis</i> (IC ₅₀ μM)	<i>G. intestinalis</i> (IC ₅₀ μM)	<i>E. histolytica</i> (IC ₅₀ μM)
Omeprazole	0.1216	0.0955	0.4922
Lansoprazole	0.1218	0.0731	0.3466
Rabeprazole	0.1057	0.0181	0.0237
Pantoprazole	0.0756	0.0157	0.0026
ABZ	1.5905	0.0370	56.33
MTZ	0.2360	1.2260	0.35

range. Particularly, pantoprazole was the most active of the PPIs studied. Based on the strong antiprotozoal activity showed, PPIs are good candidates for drug repurposing after in vivo evaluations and clinical trials. In addition, the mechanism of action of both PPIs and non 2-methylcarbamate benzimidazole derivatives as antiprotozoal agents is yet to be described and constitutes a very interesting research topic to be addressed in future research.

Acknowledgments

This study was supported by project 80093 from CONACyT. J.P.-V. acknowledges CONACyT scholarship 173896/173896 and Carso Health Institute for the PhD student scholarship S0710443. We are grateful to Amparo Tapia for carrying out the biological assays; to Rosa Isela del Villar, Nuria Estearau, Georgina Duarte, Margarita Guzmán and Marisela Gutiérrez for the analytical support and to Walter Gloria Greimel for proof reading of the manuscript.

References and notes

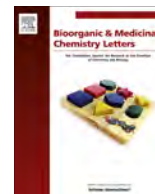
- Upcroft, P.; Upcroft, J. A. *Clin. Microbiol. Rev.* **2001**, *14*, 150.
- Berkman, D. S.; Lescano, A. G.; Gilman, R. H.; Lopez, S.; Black, M. M. *Lancet* **2002**, *359*, 564.
- Tanyuksel, M.; Petri, W. A. *Clin. Microbiol. Rev.* **2003**, *16*, 713.
- Schwebke, J. R.; Burgess, D. *Clin. Microbiol. Rev.* **2004**, *17*, 794.
- Cudmore, S. L.; Delgaty, K. L.; Hayward-McClelland, S. F.; Petrin, D. P.; Garber, G. E. *Clin. Microbiol. Rev.* **2004**, *17*, 783.
- Sears, S. D.; O'Hare, J. *Antimicrob. Agents Chemother.* **1988**, *32*, 144.
- Cedillo-Rivera, R.; Muñoz, O. *J. Med. Microbiol.* **1992**, *37*, 221.
- Chávez, B.; Cedillo-Rivera, R.; Martínez-Palomo, A. *J. Protozool.* **1992**, *39*, 510.
- Chávez, B.; Espinosa-Cantellano, M.; Cedillo Rivera, R.; Ramírez, A.; Martínez-Palomo, A. *Arch. Med. Res.* **1992**, *23*, 63.
- Navarrete-Vázquez, G.; Cedillo, R.; Hernández-Campos, A.; Yépez, L.; Hernández-Luis, F.; Valdez, J.; Morales, R.; Cortés, R.; Hernández, M.; Castillo, R. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 187.
- Navarrete-Vázquez, G.; Yépez-Mulia, L.; Cedillo-Rivera, R.; Tapia, A.; Vilpo, L.; Vilpo, J.; Kazimierczuk, Z. *Eur. J. Med. Chem.* **2002**, *37*, 973.
- Navarrete-Vázquez, G.; Yépez, L.; Hernández-Campos, A.; Tapia, A.; Hernández-Luis, F.; Cedillo, R.; González, J.; Martínez-Fernández, A.; Martínez-Grueiro, M.; Castillo, R. *Bioorg. Med. Chem.* **2003**, *11*, 4615.
- Andrzejewska, M.; Yépez-Mulia, L.; Tapia, A.; Cedillo-Rivera, R.; Laudy, A. E.; Starosciak, B. J.; Kazimierczuk, Z. *Eur. J. Pharm. Sci.* **2004**, *21*, 323.
- Navarrete-Vázquez, G.; Rojano-Vilchis, M. D.; Yépez-Mulia, L.; Meléndez, V.; Gerena, L.; Hernández-Campos, A.; Castillo, R.; Hernández-Luis, F. *Eur. J. Med. Chem.* **2006**, *41*, 135.
- Valdez-Padilla, D.; Rodríguez-Morales, S.; Hernández-Campos, A.; Hernández-Luis, F.; Yépez-Mulia, L.; Tapia-Contreras, A.; Castillo, R. *Bioorg. Med. Chem.* **2009**, *17*, 1724.
- Hernández-Luis, F.; Hernández-Campos, A.; Castillo, R.; Navarrete-Vázquez, G.; Soria-Arteche, O.; Hernández-Hernández, M.; Yépez-Mulia, L. *Eur. J. Med. Chem.* **2010**, *45*, 3135.
- López-Vallejo, F.; Medina-Franco, J. L.; Hernández-Campos, A.; Rodríguez-Morales, S.; Yépez, L.; Cedillo, R.; Castillo, R. *Bioorg. Med. Chem.* **2007**, *15*, 1117.
- Pérez-Villanueva, J.; Medina-Franco, J. L.; Caulfield, T. R.; Hernández-Campos, A.; Hernández-Luis, F.; Yépez-Mulia, L.; Castillo, R. *Eur. J. Med. Chem.* **2011**, *46*, 3499.
- Pérez-Villanueva, J.; Santos, R.; Hernández-Campos, A.; Giulianotti, M. A.; Castillo, R.; Medina-Franco, J. L. *Bioorg. Med. Chem.* **2010**, *18*, 7380.
- Pérez-Villanueva, J.; Santos, R.; Hernández-Campos, A.; Giulianotti, M. A.; Castillo, R.; Medina-Franco, J. L. *Med. Chem. Commun.* **2011**, *2*, 44.
- Boiani, M.; Gonzalez, M. *Mini-Rev. Med. Chem.* **2005**, *5*, 409.
- Horn, J. *Clin. Ther.* **2000**, *22*, 266.
- Maggiore, G. M. *J. Chem. Inf. Model.* **2006**, *46*, 1535.
- Chong, C. R.; Sullivan, D. J., Jr. *Nature* **2007**, *448*, 645.
- Conformational analysis for molecules with rotatable bonds was conducted by means of the Monte Carlo method with torsions of 30° using MMFF94 force field. Equilibrium geometry and molecular electrostatic potential calculations were carried out over selected minimum energy conformers at the Hartree-Fock level using the 6–31G^{*} basis set.
- Spartan'02, Wavefunction Inc., Irvine, CA.
- SYBYL 7.3, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.
- Klebe, G. *J. Mol. Med.* **2000**, *78*, 269.
- Omeprazole was obtained from 'Losec A', produced by AstraZeneca AB (Sweden), lansoprazole from 'Serax', produced by Landsteiner Scientific S.A. de C.V. (Mexico), rabeprazole from 'Pariet', produced by Eisai Co. Ltd. (Japan) and pantoprazole from 'Pantozol', produced by Nycomed Oranienburg GmbH (Germany). The content of the capsules equivalent to 200 mg of drug were

finely crushed in a mortar and were suspended in 20 mL of a 1% aqueous solution of sodium sulfite. All compounds were extracted with 20 mL of ethyl acetate and the suspended solids were separated by vacuum filtration using Whatman paper No. 1. The organic phase was dried with anhydrous sodium sulfate and evaporated under reduced pressure. The evaporation residue for omeprazole was crystallized using acetonitrile and the residues for lansoprazole and pantoprazole were treated with cold hexane and diisopropyl ether respectively to give white solids in all cases. Rabeprezole was isolated as sodium salt; the evaporation residue was treated with 1 equiv of NaOH 0.1 N to give an aqueous solution and then distilled under reduced pressure; finally diisopropyl ether was added and the precipitate product was filtrated under vacuum to give a white solid (rabeprezole sodium).

31. 5(6)-Methoxy-2-[[[4-methoxy-3,5-dimethylpyridin-2-yl)methyl]sulfinyl]-1H-benzimidazole (omeprazole). Acetonitrile; mp 155–157 °C (dec.); ¹H NMR (DMSO-*d*₆) δ 2.19 (s, 3H), δ 2.22 (s, 3H), δ 3.71 (s, 3H), δ 3.83 (s, 3H), 4.68–4.82 (AB system, 2d, *J* = 13.6 Hz, 2H), δ 6.95 (dd, *J* = 8.9 and 2.4 Hz, 1H), δ 7.12 (s, 1H), δ 7.57 (d, *J* = 8.9 Hz, 1H), δ 8.21 (s, 1H), δ 13.47 (s, 1H); FAB-MS *m/z*: 346 [M+1]⁺. 2-[[[3-Methyl-4-(2,2,2-trifluoroethoxy)pyridin-2-yl)methyl]sulfinyl]-

1H-benzimidazole (lansoprazole). Hexane; mp 165–167 °C (dec.); ¹H NMR (DMSO-*d*₆) δ 2.20 (s, 3H), δ 4.76–4.89 (AB system, *J* = 13.8 Hz, 2H), δ 4.94 (q, *J* = 8.7 Hz, 1H), δ 7.12 (d, *J* = 5.7 Hz, 1H), δ 7.33 (s, 2H), δ 7.62 (s, 1H), δ 7.73 (s, 1H), δ 8.32 (d, *J* = 5.7 Hz, 1H), δ 13.62 (s, 1H); FAB-MS *m/z*: 370 [M+1]⁺. Sodium 2-[[[4-(3-methoxypropoxy)-3-methylpyridin-2-yl)methyl]sulfinyl]benzimidazol-1-ide (rabeprezole). Diisopropyl ether; mp 133–135 °C (dec.); ¹H NMR (DMSO-*d*₆) δ 1.99 (m, *J* = 6.2 Hz, 2H), δ 2.17 (s, 3H), δ 3.26 (s, 3H), δ 3.50 (t, *J* = 6.2 Hz, 2H), δ 4.10 (t, *J* = 6.2 Hz, 2H), δ 4.45–4.77 (AB system, 2d, *J* = 13.0 Hz, 2H), δ 6.94–6.97 (m, 3H), δ 7.49–7.52 (m, 2H), δ 8.28 (d, *J* = 5.7 Hz, 1H); FAB-MS *m/z*: 382 [M+1]⁺. 5(6)-(Difluoromethoxy)-2-[[[3,4-dimethoxypyridin-2-yl)methyl]sulfinyl]-1H-benzimidazole (pantoprazole). Diisopropyl ether; mp 137–139 °C (dec.); ¹H NMR (DMSO-*d*₆) δ 3.80 (s, 3H), δ 3.92 (s, 3H), δ 4.67–4.77 (AB system, 2d, *J* = 13.0 Hz, 2H), δ 7.14 (d, *J* = 5.6 Hz, 1H), δ 7.28 (t, *J* = 7.4 Hz, 1H), δ 7.19 (d, *J* = 8.6 Hz, 1H), δ 7.47 (s, 1H), δ 7.71 (br, 1H), δ 8.18 (d, *J* = 5.5 Hz, 1H), δ 13.80 (s, 1H); FAB-MS *m/z*: 384 [M+1]⁺.

32. Ashburn, T. T.; Thor, K. B. *Nat. Rev. Drug Disc.* **2004**, *3*, 673.
 33. O'Connor, K. A.; Roth, B. L. *Nat. Rev. Drug Disc.* **2005**, *4*, 1005.
 34. Dolgin, E. *Nat. Med.* **2011**, *17*, 1027.



Synthesis and antiprotozoal activity of novel 2-[[2-(1*H*-imidazol-1-yl)ethyl]sulfonyl]-1*H*-benzimidazole derivatives

Jaime Pérez-Villanueva^a, Alicia Hernández-Campos^a, Lilián Yépez-Mulia^b, Carlos Méndez-Cuesta^a, Oscar Méndez-Lucio^a, Francisco Hernández-Luis^a, Rafael Castillo^{a,*}

^a Facultad de Química, Departamento de Farmacia, UNAM, DF 04510, Mexico

^b Unidad de Investigación Médica en Enfermedades Infecciosas y Parasitarias, IMSS, DF 06720, Mexico

ARTICLE INFO

Article history:

Received 18 February 2013

Revised 1 May 2013

Accepted 7 May 2013

Available online 15 May 2013

Keywords:

Benzimidazole

Entamoeba histolytica

Giardia intestinalis

Trichomonas vaginalis

ABSTRACT

A series of 19 new 2-[[2-(1*H*-imidazol-1-yl)ethyl]sulfonyl]-1*H*-benzimidazole derivatives was synthesized starting from the properly substituted 1,2-phenyldiamine. These compounds have hydrogen or methyl at position 1; while hydrogen, chlorine, ethoxy or methoxycarbonyl group is at position 5 and/or 6. The novel compounds were tested against protozoa *Trichomonas vaginalis*, *Giardia intestinalis* and *Entamoeba histolytica*. Experimental evaluations revealed strong activity for all tested compounds, having IC₅₀ values in the nanomolar range, which were even better than metronidazole, the drug of choice for these parasites.

© 2013 Elsevier Ltd. All rights reserved.

Parasitic infections caused by protozoa still represent a major public health problem in developing countries. Some important intestinal protozoa include *Giardia intestinalis* and *Entamoeba histolytica*, causal agents of giardiasis and amebiasis, respectively. According to the World Health Organization (WHO), there are estimated 280 million giardiasis cases each year. Interestingly, *G. intestinalis* has been classified as the most common diagnosed flagellate in the intestinal tract.^{1,2} Infection by this protozoan usually produces diarrhea and associated symptoms. Furthermore, it is worth mentioning that child incidence is very high and is associated with malnutrition, linear growth retardation and poor cognitive function.^{3,4} Regarding *E. histolytica*, WHO estimated that this protozoan causes severe disease in 50 million people each year and has been classified as one of the most common causes of death from parasitic disease.^{2,5} In addition to the common symptoms such as diarrhea and dysentery, this protozoan can penetrate the intestinal mucosa and migrate to other organs causing severe damage.^{1,2,5} In addition to intestinal infections, the genitourinary infection caused by *T. vaginalis* (trichomonosis) is estimated to be more than 180 million new cases annually.¹ Although trichomonosis in men is generally asymptomatic or mild, in women it causes severe symptoms and consequences. The infection produces deep inflammation of the genital tract and has been associated with preterm labor, low-birth weight, sterility, cervical cancer and a predisposition to HIV infection.^{6,7} For these three diseases, metronidazole (MTZ)

has been successfully used as the drug of choice for more than 40 years; however, its side effects and the development of resistant strains limit its use.¹ Although some additional chemotherapeutic agents are available (e.g. tinidazole and nitazoxanide used in the treatment of giardiasis), it is still important to have more options of treatment, because of different individual response to drugs. During the last years, an important number of benzimidazole derivatives have been synthesized and evaluated as antiprotozoals by our group and other external research groups.^{8–15} Moreover, studies based on the emerging concept of the activity landscape were undertaken to find out the structure–activity relationships (SAR) of benzimidazole derivatives as trichomonocidal and giardicidal agents.^{16,17} Besides, quantitative structure–activity relationships (QSAR) studies based on comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were carried out for amebicidal and trichomonocidal benzimidazole derivatives.^{18,19} Although some important conclusions have been highlighted with our previous synthetic and SAR studies, the information accumulated is still limited. Therefore, increasing the database of benzimidazole derivatives remains of paramount importance to have access to new SAR features that can lead to the optimization of benzimidazole derivatives as antiprotozoals. Expanding our previous work, herein we report the synthesis of 19 new 2-[[2-(1*H*-imidazol-1-yl)ethyl]sulfonyl]-1*H*-benzimidazole derivatives (**35–53**), Fig. 1.

The design of the new benzimidazole derivatives was based on previous SAR conclusions derived from our former studies (Fig. 1).^{16–19} It was previously observed that substituted alkylthio

* Corresponding author. Tel.: +52 556 22 52 87; fax: +52 556 22 53 29.
E-mail address: rafaelc@unam.mx (R. Castillo).

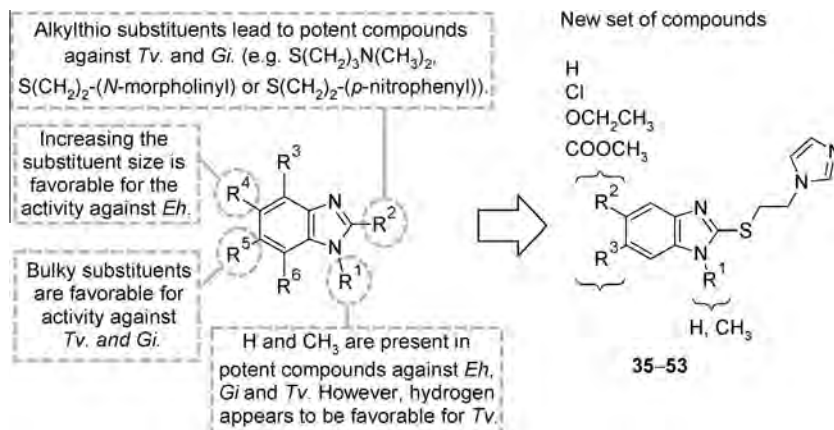


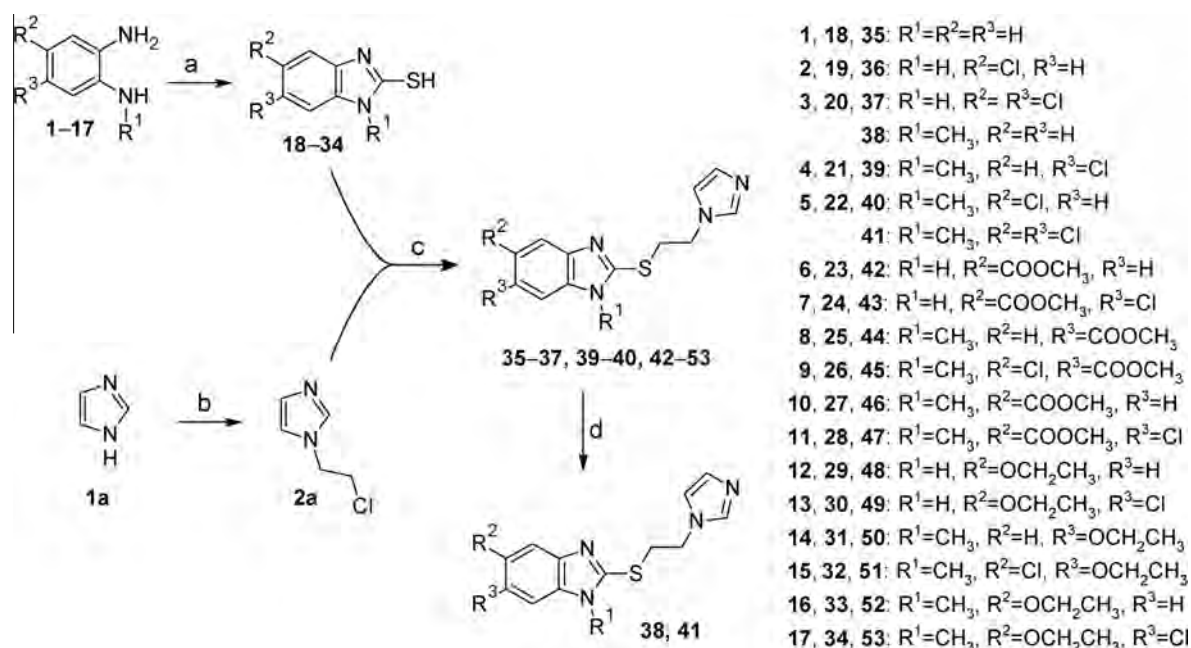
Figure 1. Some SAR conclusions for benzimidazole derivatives tested against *G. intestinalis* (*Gi*), *E. histolytica* (*Eh*), and *T. vaginalis* (*Tv*) which lead to the structure of the 19 new derivatives (**35–53**) synthesized in this work.

groups at position 2 of the benzimidazole nucleus lead to potent compounds against *T. vaginalis* and *G. intestinalis*.^{12,17,19} Some of the most active compounds were those with hydrogen acceptors or an aromatic ring attached to the alkyl chain. The new derivatives synthesized in this work contain an alkylthio imidazole group at position 2, a substituent which combines both features. In addition, hydrogen or methyl substituents were present at position 1 of the new derivatives. It is worth noting that both, hydrogen and methyl derivatives, have been reported as strong trichomonocidal and giardicidal agents.^{9,14} Also, hydrogen, chlorine, carbomethoxy and ethoxy substituents were considered at position 5 and/or 6 for the new derivatives. As previous QSAR studies of benzimidazole derivatives suggest, the substituents at position 6 are favorable for trichomonocidal and giardicidal activity,^{16,17,19} whereas substituents at position 5 play an important role for amebicidal activity.¹⁸ It is important to emphasize that SAR conclusions previously reported are highly dependent on the database analyzed.

Compounds **35–53** were prepared according to the sequence of reactions shown in Scheme 1. The first step was the cyclocondensation of the properly substituted 1,2-phenylenediamine (**1–17**)

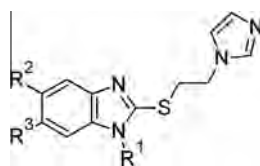
with carbon disulfide in the presence of potassium hydroxide and ethanol to give the respective 2-mercaptobenzimidazole (**18–34**).^{20,21} These were then alkylated with 1-(2-chloroethyl)-1*H*-imidazole to afford title compounds **35–37**, **39**, **40**, and **42–53**.²² Finally, compounds **38** and **41** were obtained by *N*-methylation of compounds **35** and **37**, respectively.²³ The intermediate 1-(2-chloroethyl)-1*H*-imidazole was prepared under solid–liquid phase transfer-catalysis conditions as described previously by Bogdal et al.,²⁴ in which commercial imidazole was alkylated with 1,2-dichloroethane in the presence of sodium hydroxide, potassium carbonate and tetrabutylammonium chloride. Compounds **35–53** were obtained as solids with fair global yields and purity. The compounds were characterized using ¹H and ¹³C nuclear magnetic resonance,²⁵ mass spectrometry and high resolution mass spectrometry (See supplementary data for structure characterization, yields and melting points).

The in vitro antiprotozoal activity of compounds **35–53** was tested following the method previously described for *T. vaginalis*, *G. intestinalis* and *E. histolytica*.^{8–15,26} The results obtained from the biological assays are shown in Table 1. All tested compounds



Scheme 1. Synthesis of compounds **35–53**: (a) CS_2 , KOH, EtOH/ H_2O , 40–50 °C, 8 h; (b) $ClCH_2CH_2Cl$, Bu_4NCl , NaOH, K_2CO_3 , 50 °C, 3 h; (c) CH_3CN , KOH or DBU, 60 °C, 8 h; (d) CH_3I , $(CH_3)_2CO$, K_2CO_3 .

Table 1
Chemical structures and biological activity of benzimidazole derivatives **35–53**



	R ₁	R ₂	R ₃	IC ₅₀ (μM) <i>T. vaginalis</i>	IC ₅₀ (μM) <i>G. intestinalis</i>	IC ₅₀ (μM) <i>E. histolytica</i>
35	H	H	H	0.1780 ± 0.0087	0.1167 ± 0.0029	0.1474 ± 0.0058
36	H	Cl	H	0.1704 ± 0.0076	0.0807 ± 0.0025	0.0735 ± 0.0076
37	H	Cl	Cl	0.1405 ± 0.0045	0.0766 ± 0.0090	0.0591 ± 0.0023
38	CH ₃	H	H	0.1490 ± 0.0027	0.0871 ± 0.0082	0.0774 ± 0.0055
39	CH ₃	H	Cl	0.1281 ± 0.0072	0.0626 ± 0.0039	0.0672 ± 0.0110
40	CH ₃	Cl	H	0.1349 ± 0.0024	0.0666 ± 0.0024	0.0660 ± 0.0129
41	CH ₃	Cl	Cl	0.0963 ± 0.0022	0.0688 ± 0.0065	0.0642 ± 0.0043
42	H	COOCH ₃	H	0.1108 ± 0.0023	0.0777 ± 0.0023	0.0612 ± 0.0023
43	H	COOCH ₃	Cl	0.0727 ± 0.0105	0.0549 ± 0.0021	0.0579 ± 0.0021
44	CH ₃	H	COOCH ₃	0.1138 ± 0.0055	0.0569 ± 0.0045	0.0522 ± 0.0067
45	CH ₃	Cl	COOCH ₃	0.0869 ± 0.0101	0.0556 ± 0.0020	0.0428 ± 0.0040
46	CH ₃	COOCH ₃	H	0.0774 ± 0.0067	0.0616 ± 0.0067	0.0427 ± 0.0112
47	CH ₃	COOCH ₃	Cl	0.0926 ± 0.0020	0.0442 ± 0.0020	0.0513 ± 0.0040
48	H	OCH ₂ CH ₃	H	0.0717 ± 0.0053	0.0566 ± 0.0020	0.0428 ± 0.0044
49	H	OCH ₂ CH ₃	Cl	0.0991 ± 0.0044	0.0465 ± 0.0044	0.0620 ± 0.0044
50	CH ₃	H	OCH ₂ CH ₃	0.0717 ± 0.0038	0.0452 ± 0.0083	0.0408 ± 0.0019
51	CH ₃	Cl	OCH ₂ CH ₃	0.0698 ± 0.0063	0.0356 ± 0.0084	0.0460 ± 0.0063
52	CH ₃	OCH ₂ CH ₃	H	0.0761 ± 0.0094	0.0083 ± 0.0023	0.0298 ± 0.0047
53	CH ₃	OCH ₂ CH ₃	Cl	0.0980 ± 0.0042	0.0208 ± 0.0042	0.0148 ± 0.0042
Metronidazole				0.2360 ± 0.0160	1.2260 ± 0.1250	0.3798 ± 0.1461
Albendazole				1.5905 ± 0.0113	0.0370 ± 0.0030	56.5334 ± 18.8445

showed good activity in the nanomolar range. It is worth noting that all compounds were more potent than metronidazole, the drug of choice for these three parasites. The SAR nature of the database presented in this work is essentially continuous. That is, the new set of compounds is in agreement with the similarity principle which establishes that structurally similar compounds have similar biological activities. This conclusion is supported since all molecules in the dataset have less than 15-fold variation in potency. However, these low variations are enough to support some SAR conclusions and are in agreement with most of the observations previously reviewed (Fig. 1).

Results in Table 1, for compounds **35–41**, show that dichloro substitution at positions 5 and 6 in compounds **37** and **41** is slightly favorable for antiprotozoal activity, as compared with non-substituted compounds **35** and **38**, respectively. Furthermore, compounds **35** and **38** show a favorable effect of the methyl group at position 1 against the three protozoa tested. This same tendency was observed for compounds **37** and **41** against *T. vaginalis*, but no important changes were observed for *G. intestinalis* and *E. histolytica*. It is interesting to note that all the 1-methyl compounds show comparable or better activity than their respective 1H analogue. The only exceptions are compounds **44** and **47**, which have a slightly lower activity against *T. vaginalis* as compared with their 1H analogues **42** and **43**, respectively. These results agree with the previous SAR models that suggest that hydrogen at position 1 is favorable for trichomonocidal activity. In general, compounds **43–53**, which have a carbomethoxy or an ethoxy substituent at positions 5, 6 present the best activities. In particular, compounds **50–53**, with an ethoxy substituent, have similar or higher giardicidal potency than albendazole. It is important to emphasize that compound **52**, which showed a very strong giardicidal activity, was the only one whose activity increased 14-fold against *G. intestinalis*, as compared with unsubstituted analogue **35**, which was the less active compound against all the protozoa tested. Additionally, compounds **50–53** were particularly active against *T. vaginalis* and *E. histolytica* as compared with albendazole, which is poorly

active. Compounds **50–53** can be considered among the most potent benzimidazole derivatives reported against the three protozoa. Noteworthy, antiprotozoal activities are in agreement with the previous SAR studies that indicate the importance of the substituent size at position 5 and 6 of the benzimidazole nucleus for the antiprotozoal activity.^{16–19} Also, the strong activity of compounds **35–53** agrees with the previous SAR observation that alkylthio groups at position 2, having a hydrogen acceptor or an aromatic ring, lead to potent compounds.

The mechanism of action for the antiprotozoal activity of compounds **35–53** is still unknown. Interestingly, we found in previous studies that benzimidazole derivatives that are not benzimidazole carbamates (BC) do not inhibit β-tubulin polymerization as anti-parasitic BC do.^{8,9} Therefore, a different mechanism of action is implied in the antiprotozoal activity of the current derivatives.

In summary, compounds **35–53** resulted in new structures which were able to be synthesized with acceptable yields starting from the properly substituted 1,2-phenyldiamine. The novel compounds were characterized by using spectrometric and spectroscopic methods. Biological assays revealed that the novel compounds have strong activity against *T. vaginalis*, *G. intestinalis* and *E. histolytica*, which is better than metronidazole. Particularly, compounds **50–53** exhibited the best activities; among these, compound **52** was the most active against *G. intestinalis*. The new compounds expand our database of benzimidazole derivatives and increase our knowledge concerning the structural requirements for antiprotozoal activity. These results are useful for the design of new stronger antiprotozoal compounds.

Acknowledgements

This work was supported by project 80093 from CONACyT. J.P.-V. acknowledges the CONACyT scholarship 173896/173896 and Carso Health Institute for the PhD student scholarship S0710443. We are grateful to Amparo Tapia for carrying out the biological assays and to Rosa Isela del Villar, Nuria Estearau,

Georgina Duarte, Margarita Guzmán and Marisela Gutiérrez for the analytical support.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bmcl.2013.05.012>.

References and notes

1. Upcroft, P.; Upcroft, J. A. *Clin. Microbiol. Rev.* **2001**, *14*, 150.
2. Marshall, M. M.; Naumovitz, D.; Ortega, Y.; Sterling, C. R. *Clin. Microbiol. Rev.* **1997**, *10*, 67.
3. Guerrant, R. L.; Schorling, J. B.; McAuliffe, J. F.; Desouza, M. A. *Am. J. Trop. Med. Hyg.* **1992**, *47*, 28.
4. Berkman, D. S.; Lescano, A. G.; Gilman, R. H.; Lopez, S.; Black, M. M. *Lancet* **2002**, *359*, 564.
5. Tanyuksel, M.; Petri, W. A. *Clin. Microbiol. Rev.* **2003**, *16*, 713.
6. Schwabke, J. R. *Sex. Transm. Infect.* **2002**, *78*, 378.
7. Cudmore, S. L.; Delgaty, K. L.; Hayward-McClelland, S. F.; Petrin, D. P.; Garber, G. E. *Clin. Microbiol. Rev.* **2004**, *17*, 783.
8. Navarrete-Vázquez, G.; Cedillo, R.; Hernández-Campos, A.; Yépez, L.; Hernández-Luis, F.; Valdez, J.; Morales, R.; Cortés, R.; Hernández, M.; Castillo, R. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 187.
9. Valdez, J.; Cedillo, R.; Hernández-Campos, A.; Yépez, L.; Hernández-Luis, F.; Navarrete-Vázquez, G.; Tapia, A.; Cortés, R.; Hernández, M.; Castillo, R. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2221.
10. Navarrete-Vázquez, G.; Yépez-Mulia, L.; Cedillo-Rivera, R.; Tapia, A.; Vilpo, L.; Vilpo, J.; Kazimierczuk, Z. *Eur. J. Med. Chem.* **2002**, *37*, 973.
11. Navarrete-Vázquez, G.; Yépez, L.; Hernández-Campos, A.; Tapia, A.; Hernández-Luis, F.; Cedillo, R.; González, J.; Martínez-Fernández, A.; Martínez-Grueiro, M.; Castillo, R. *Bioorg. Med. Chem.* **2003**, *11*, 4615.
12. Andrzejewska, M.; Yépez-Mulia, L.; Tapia, A.; Cedillo-Rivera, R.; Laudy, A. E.; Starosciak, B. J.; Kazimierczuk, Z. *Eur. J. Pharm. Sci.* **2004**, *21*, 323.
13. Navarrete-Vázquez, G.; Rojano-Vilchis, M. D.; Yépez-Mulia, L.; Meléndez, V.; Gerena, L.; Hernández-Campos, A.; Castillo, R.; Hernández-Luis, F. *Eur. J. Med. Chem.* **2006**, *41*, 135.
14. Valdez-Padilla, D.; Rodríguez-Morales, S.; Hernández-Campos, A.; Hernández-Luis, F.; Yépez-Mulia, L.; Tapia-Contreras, A.; Castillo, R. *Bioorg. Med. Chem.* **2009**, *17*, 1724.
15. Hernández-Luis, F.; Hernández-Campos, A.; Castillo, R.; Navarrete-Vázquez, G.; Soria-Arteche, O.; Hernández-Hernández, M.; Yépez-Mulia, L. *Eur. J. Med. Chem.* **2010**, *45*, 3135.
16. Pérez-Villanueva, J.; Santos, R.; Hernández-Campos, A.; Giulianotti, M. A.; Castillo, R.; Medina-Franco, J. L. *Bioorg. Med. Chem.* **2010**, *18*, 7380.
17. Pérez-Villanueva, J.; Santos, R.; Hernández-Campos, A.; Giulianotti, M. A.; Castillo, R.; Medina-Franco, J. L. *Med. Chem. Commun.* **2011**, *2*, 44.
18. López-Vallejo, F.; Medina-Franco, J. L.; Hernández-Campos, A.; Rodríguez-Morales, S.; Yépez, L.; Cedillo, R.; Castillo, R. *Bioorg. Med. Chem.* **2007**, *15*, 1117.
19. Pérez-Villanueva, J.; Medina-Franco, J. L.; Caulfield, T. R.; Hernández-Campos, A.; Hernández-Luis, F.; Yépez-Mulia, L.; Castillo, R. *Eur. J. Med. Chem.* **2011**, *46*, 3499.
20. VanAllan, J. A.; Deacon, B. D. *Org. Synth. Coll. Vol.* **1963**, *4*, 569.
21. General procedure for the synthesis of 2-mercaptobenzimidazole derivatives (**18–34**): the proper 1,2-phenylenediamine **1–17** (16 mmol) was added over a stirred solution of ethanol (50 mL), water (5 mL), potassium hydroxide (19.2 mmol) and carbon disulfide (19.2 mmol). The mixture was heated overnight under a N₂ atmosphere at 50 °C. The cold reaction mixture was then neutralized with acetic acid and poured into 250 mL of cold water for complete precipitation. The solid was separated and dried using vacuum filtration. The crude product was purified by recrystallization, adding norite to the solvent.
22. General procedure for the synthesis of 2-[(2-(1H-imidazol-1-yl)ethyl)sulfanyl]-1H-benzimidazole derivatives (**35–37**, **39**, **40** and **42–53**): the proper 2-mercaptobenzimidazole **18–34** (6 mmol) in 20 mL of acetonitrile was stirred at room temperature and treated with KOH (6.6 mmol, as a 50% aqueous solution) or DBU (6.6 mmol, for ester derivatives **23–28**). Then, 1-(2-chloroethyl)-1H-imidazole (10.2 mmol) was added and the mixture was heated at 60 °C for 6–8 h. Next, the cold reaction mixture was neutralized with acetic acid and the solvent was removed under vacuum. The evaporation residue was extracted with ethyl acetate (30–50 mL), the organic phase was washed thrice with 20 mL of water, then dried with anhydrous sodium sulfate, and evaporated under vacuum. The residue was purified by column chromatography using ethyl acetate as a mobile phase.
23. General procedure for the synthesis of 2-[(2-(1H-imidazol-1-yl)ethyl)sulfanyl]-1H-benzimidazole derivatives (**38** and **41**): a solution of the proper benzimidazole derivative, **35** or **37** (2 mmol) in 2-propanone (5 mL) was treated with KOH (2.2 mmol, as a 50% aqueous solution). Then, methyl iodide (2.2 mmol) was added and the mixture was stirred for 4 h at room temperature. Finally, the reaction was poured into 50 mL of cold water and the solid product was collected on a Büchner funnel. The crude product was purified by filtration using a column packed with silica gel using ethyl acetate as solvent.
24. Bogdal, D.; Jaskot, K. *Synth. Commun.* **2000**, *30*, 3341.
25. Some signals in ¹³C NMR for 1H substituted benzimidazole derivatives are duplicated due to tautomeric effect. This same effect decreases the intensity of ¹³C signals; therefore, in some compounds signals are weak or can be missed.
26. Biological assays: *Trichomonas vaginalis* strain GT3, *Giardia intestinalis* isolate IMSS:0981:1 and *Entamoeba histolytica* strain HM1-IMSS were used in all the experiments. Trophozoites of *G. intestinalis* were maintained in a TYI-S-33 medium supplemented with 10% calf serum and bovine bile. *E. histolytica* and *T. vaginalis* trophozoites were maintained in TYI-S-33 medium supplemented with 10% bovine serum. Briefly, 5 × 10⁴ trophozoites of *G. intestinalis* or *T. vaginalis*, or 6 × 10³ trophozoites of *E. histolytica* were incubated for 48 h at 37 °C with different concentrations of the compound to be tested, each added as solutions in DMSO. As a negative control, parasite cultures received an equivalent amount of DMSO only, while ABZ and MTZ were included as positive controls. At the end of the treatment period, the cells were washed and subcultured for another 48 h in a fresh medium to which no drug was added. The trophozoites were then counted with a haemocytometer and the 50% inhibitory concentration (IC₅₀), together with the respective 95% confidence limit was calculated by Probit analysis. Experiments were carried out in triplicate and repeated at least twice.