Electronic Supplementary Material (ESI) for RSC Advances.

Elucidating Arsenic-Bound Proteins in the Protein Data Bank: Data Mining and Amino Acid Cross-Validation through Raman Spectroscopy

Upendra Nayek[1], Sudarshan Acharya[1], Abdul Ajees Abdul Salam[1*]

[1]Department of Atomic and Molecular Physics, Manipal Academy of Higher Education, Manipal, Karnataka 576 104, India.

[*]Corresponding Author: abdul.ajees@manipal.edu; Phone: +91 8147966458

Supplementary Information

Supplementary Table 1 Exploring arsenical protein ids and specified arsenical ligands [As(III)] through data mining: unveiling insights and patterns.

Ligand ID	Ligand Name	Chemical formulas	Molecular weight	Total No. (902)	PDB ID
RMO	[arsenothionito(2-)-kappa~2~O,S](oxo)molybdenum	$AsHMoO_3S$	251.93	1	3SR6
RXO	4-arsanyl-2-nitrophenol	$C_6H_6AsNO_3$	215.04	2	4RSR, 5V0F
5AU	Di-glutathione-PhenylArsine	$C_{26}H_{37}AsN_6O_{12}S_2$	764.66	2	5DAL, 5DDL
CZ2	Sdihydroxyarsino)cysteine	$C_3H_8AsNO_4S$	229.09	3	1SK0, 1SJZ, 6EU7
CZZ	Thiarsahydroxy-	C_3H_7AsN	212.08	5	1J9B, 6YH0, 6TSQ, 6TSR,

Ligand ID	Ligand Name	Chemical formulas	Molecular weight	Total No. (902)	PDB ID
	cysteine	O_3S			6TSN
PA0	Phenylarsine oxide	C_6H_5AsO	168.03	5	3E3Z, 4KW7, 5EG5, 5DAL, 6QHK
TAS/ AST	Trihydroxyarsenite(iii)	AsH_3O_3	125.94	11	1II0, 1IHU,1II9, 1SJZ, 1SK0, 1J9B, 6DUN, 1SIJ, 3NVV, 3L4P, 6CZ9
ARS	Arsenic	As	74.92	49	1F0J,1NLX,1NQ0,1NQ1,1NQ2,1RO6,1WN6,1Y0R,2GBM,2IA5,2J4Z,2JGO,2O6W,2XQ4,2XQ5,2XQ9,2Y0O,3CAO,3CAR,3CFS,3CFV,3P,3FRG,3GWT,3H6N,3JRN,3O56,3O57,3SMP,3SMT,4DMN,4FSD,4GNJ,4GVM,4GW6,4J9T,4R0V,5D8Q,5DAK,6CX6,6J05,6J0E,6L0C,6UDN,7DHY,7DHZ,7EDS,7EEU,7V97
CAS	S-(dimethyl arsenic)cysteine	$C_5H_{12}AsNO_2S$	225.14	228	1BHL,1FF3,1FU1,1FYW,1FYX,1I83,1JQ6,1R4X,1RA6,1RA7,1S6C,1TQL,1ZZS,1ZZT,2CJW,2H77,2H79,2HMH,2HZE,2ILY,2ILZ,2IM0,2IM1,2IM2,2IM3,2JBH,2RA8,2W9J,3E7S,3ET6,3G2F,3GWS,3HZF,3ILZ,3IMY,3JZB,3JZC,3P4U,3PSZ,4AWJ,4B09,4B9K,4BKS,4BKT,4C3A,4CAR,4CFT,4CTY,4CTZ,4CU0,4CU1,4CUL,4CUM,4CUN,4CVG,4CWV,4CWW,4CWX,4CWY,4CWZ,4CX0,4CX1,4CX2,4D34,4D35,4D36,4D37,4D38,4D39,4D3A,4EG1,4EG3,4EG4,4EG6,4EG7,4EG8,4EGA,4IKB,4JQW,4JSK,4JSL,4JSM,4K5H,4K5I,4K5J,4K5K,4KCP,4KCQ,4KCR,4KCS,4L3T,4LH4,4LH5,4LNW,4LNX,4LUW,4MVW,4MVX,4MVY,4MW0,4MW1,4MW2,4MW4,4MW5,4MW6,4MW7,4MWB,4MWC,4MWD,4MWE,4NGE,4NLO,4NLP,4NLQ,4NLR,4NLS,4NLT,4NLU,4NLV,4NLW,4NLX,4NLY,4RPU,4UH7,4UH8,4UH9,4UHA,4UPQ,4UPR,4UPS,4UPT,4W9C,4W9D,4W9E,4W9F,4W9G,4W9H,4W9I,4W9J,4W9K,4W9L,4WP2,4XA5,4ZT2,4ZT3,4ZT4,4ZT5,4ZT6,4ZT7,5ADJ,5ADK,5ADL,5ADM,5ADN,5BO4,5BYK,5CG7,5DBX,5FJ2,5FJ3,5FVY,5FVZ,5GPG,5HR7,5LLI,5N6N,5NAQ,5NVV,5NVW,5NVX,5NVY,5NVZ,5NW0,5NW1,5NW2,5OI2,5OI3,5OI5,5OI8,5OIA,5UOD,5VV6,5VV7,5VV8,5VV9,5VVA,5VVG,5VVN,5YV8,5YV9,5YVA,5YVB,5YVC,5ZM5,6B4X,6B4Z,6BDP,6BDR,6E66,6EXU,6FMI,6FMJ,6F

Ligand ID	Ligand Name	Chemical formulas	Molecular weight	Total No. (902)	PDB ID
					MK,6GFY,6GFZ,6GMN,6GMQ,6GMX,6I4X,6I5J,6I5N,6J6X,6MFE,6UNP,6UUX,6VPH,7A9W,7A9X,7CYS,7M6T,7PVN,8BDI,8BDJ,8BDL,8BDM,8BDN,8BDO,8BV2

Supplementary Table 2 Exploring arsenical protein ids and specified arsenical ligands [As(V)] through data mining: unveiling insights and patterns.

Ligand ID	Ligand Name	Chemical formulas	Molecular weight	Total No. (902)	PDB ID
TTA	Tetraphenyl-arsonium	$C_{24}H_{20}As$	383.34	1	1HYV
ATS	γ-arsono-β, γ-methyleneadenosine-5'-diphosphate	$C_{11}H_{18}AsN_5O_{12}P_2$	549.16	1	1GLJ
LVQ	Octa-anionic calixarene	$C_{38}H_{37}AsNaO_{25}S_4$	1,119.86	1	6SUV
1KH	Arsenoribose	$C_5H_{13}AsO_8$	276.07	1	4JD0
ASR	4-aminophenylarsonic acid	$C_6H_8AsNO_3$	217.05	1	1N4F
GK8	{[(2S)-2-amino-3-(1H-imidazol-5-yl)propyl]oxy}(trihydroxy)-lambda~5~-	$C_6H_{13}AsN_3O_4$	266.11	1	4GK8

Ligand ID	Ligand Name	Chemical formulas	Molecular weight	Total No. (902)	PDB ID
	arsanyl				
TTO	(3,4-dihydroxy-phenyl)-triphenyl-arsonium	$C_{24}H_{20}AsO_2$	415.34	1	1HYZ
BLJ	(2S)-2-amino-4-[hydroxy(methyl) arsoryl]butanoic acid	$C_5H_{12}AsNO_4$	225.07	1	5WPH
3Q7	(trimethylarsonio) acetate	$C_5H_{11}AsO_2$	178.06	1	5NXX
T1A	Tetraethylarsonium ion	$C_8H_{20}As$	191.17	1	2BOC
1Y8	2-(trimethyl-lambda~5~-arsanyl)ethanol	$C_5H_{14}AsO$	165.09	2	4LLH, 5NXY
A6R	Arsenoplatin-1	$C_4H_8AsN_2O_4Pt$	418.12	4	5NJ1, 7BD7, 5NJ7, 6ZS8
CFQ	1-(2-nitrophenyl)-2,2,2-trifluoroethyl]-arsenocholine	$C_{13}H_{19}AsF_3NO_3$	369.21	3	2V96, 2V97, 2V98
ART/ 8AR	Arsenate	AsO_4	138.92	7	1TA4, 3ENZ, 3WE3, 6CZ8, 6XL2, 4F18, 4F19
CSR	S-arsonocysteine	$C_3H_8AsNO_5S$	245.09	4	1OKG, 1SK1, 1LJU, 1JZW
CAF	S-dimethylarsinoyl-cysteine	$C_5H_{12}AsNO_3S$	241.14	48	1HYV,1HYZ,1SNG,2C7V,2DYB,2FAU,2IWR,2WTA,2XNQ,3F7D,3LPT,3LPU,4CI1,4E1M,4E1N,4H5U,4ID1,4J8M,4JLH,4O0J,4O55,4O5B,4TSX,5KGW,5KGX,5KRS,5KRT,5MDI,6EB1,6EB2,6LMI,6LMQ,6NNA,6NUJ,6QA0,6SY0,6SYM,6V98,6XGU,6XGV,6XHA,6XHB,7CYZ,7D83,7KE0,7WCE,8CT5,8CT7

Ligand ID	Ligand Name	Chemical formulas	Molecular weight	Total No. (902)	PDB ID
CAC/ CAD	Cacodylate ion	$C_2H_6AsO_2$	136.99	518	1D0C,1D0O,1D1V,1D1W,1D1X,1D1Y,1ED4,1ED5,1ED6,3JWW,3JWX,3JWY,3JWZ,3N5P,3N5Q,3N5R,3N5S,3N5T,3N67,3N68,3N69,3N6A,3N6B,3N6C,3N6D,3N6E,3N6F,3N6G,3NLD,3NLE,3NLF,3NLG,3NLH,3NLI,3NLT,3NLU,3PNH,3RQO,3RQP,3S2S,3TO7,3TO9,4IMX,4UHF,4UHH,4ZDQ,5NSE,5VV8,5WHI,5YGE,6NSE,7CW1,7CY3,7CY9,7CZ0,7EA4,7MDP,7NSE,7SXP,8NSE,9NSE,1B92,1B9D,1B9F,1BEH,1BIZ,1C82,1D1S,1D1T,1D61,1DAN,1DM6,1DM7,1DM8,1DMI,1DMJ,1DMK,1DVA,1EQZ,1F35,1FE8,1FOI,1FOJ,1FOL,1FOO,1FOP,1GQ8,1GR0,1I19,1I9S,1I9T,1ITG,1JOD,1K2O,1KZM,1L0I,1L1D,1LZK,1MWQ,1N0Y,1NSE,1NW2,1OAR,1OB1,1P13,1P6L,1P6M,1P6N,1PQU,1Q2O,1QG3,1QH3,1QQJ,1R4V,1RS8,1RS9,1TAD,1TND,1TR7,1TYE,1TZA,1VHD,1VHO,1W0Y,1W2K,1WMB,1WN5,1WY2,1X1T,1XSL,1XZX,1Y0X,1Y9A,1YHC,1Z6B,1ZV8,2A2T,2ACR,2AEI,2AVJ,2B0P,2BO2,2BOQ,2BOU,2BOX,2BVT,2BVY,2D28,2.00E+11,2G6O,2GWN,2GWS,2H5K,2HX2,2HYP,2I0K,2IUC,2IYO,2J13,2J6E,2JFS,2NSE,2O4M,2O4Q,2OUI,2PQ3,2RL7,2V5C,2VDK,2VDL,2VEQ,2VQG,2VQH,2VQK,2VQL,2VS0,2VW8,2WA7,2X2O,2X2P,2X2Q,2XFV,2XOD,2XOE,2XZ1,2Z3I,2Z3J,3A6E,3A6F,3AAL,3BFD,3BFT,3BWW,3CS2,QS,QT,UE,3.00E+38,3ERP,3FCU,3FJU,3FKG,3FM4,3FMU,3FPC,3FPL,3G3S,3GH3,3GK1,3GK2,3GSK,3H6T,3HBZ,3HTW,3I3O,3I7D,3IUW,3KP2,3KP6,3KZP,3LLM,3M1R,3MDO,3MSZ,3N5F,3NRF,3NSE,3OD5,3OUU,3P85,3PB6,3PD8,3PQS,3RF3,3RHG,3S70,3SSG,3UCY,3UJP,3VFJ,3VP5,3W86,3W87,3W88,3WAB,3WC5,3WGH,3ZPG,4C2A,4D7L,4GY1,4H5U,4HG3,4HG5,4HGD,4HX2,4ICR,4IIS,4J5G,4K4G,4K4I,4KVH,4LCZ,4MOU,4N07,4N4R,4NSE,4OJR,4OSG,4PBE,4PBF,4PCP,4PI2,4PIV,4PTX,4RLR,4RYN,4U2V,4UFC,4UM8,4V2X,4W7U,4WB0,4WXB,4WY0,4X0L,4XAF,4XAG,4XAY,4XD3,4XD4,4XD5,4XD6,4XOE,4XWF,4YN5,4Z3N,4Z3P,4ZMA,4ZNG,5AG0,5AG1,5AOS,5AOT,5CB6,5CNX,5.00E+29,5.00E+64,5.00E+66,5.00E+93,5ED4,5EZ3,5GM3,5H8S,5HLO,5HRN,5HRP,5HRR

Ligand ID	Ligand Name	Chemical formulas	Molecular weight	Total No. (902)	PDB ID
					,5HRS,5HWA,5HXD,5IC5,5J39,5JH9,5K6Z,5KEK,5KEO,5LMC,5LTA,5M11,5NT5,5O3X,5O4E,5O8X,5OC4,5OYJ,5T5K,5UW3,5UW5,5VG2,5VT3,5W0R,5W6B,5WCP,5WCQ,5WCR,5WCW,5WIZ,5WJ0,5WMS,5XKO,5XRS,5XRT,5XTU,5ZEQ,5ZVA,5ZVB,6AML,6B4W,6BH7,6BHL,6CXZ,6D8P,6EZV,6GBL,6H2P,6HMF,6JCF,6JCG,6JP4,6KBL,6M0U,6N6Q,6PI4,6QM3,6QPH,6RCQ,6RCX,6SA0,6SJD,6U40,6UNW,6V6R,6WQG,6WR3,6WR5,6WR9,6WRB,6WRI,6WSN,6WSO,6WSP,6WSQ,6WSS,6WST,6WSU,6WSW,6WSY,6WSZ,6X6D,6X8B,6X8C,6X91,6XCO,6XDV,6XDW,6XDX,6XDZ,6XEI,6XEJ,6XEL,6XEM,6XFC,6XFD,6XFE,6XFG,6XFW,6XFY,6XFZ,6XG0,6XGJ,6XGK,6XGL,6XNA,6XO5,6XO6,6XO7,6XO9,6XU0,6XUP,6YAB,6YPE,6YPX,6Z3X,6ZYU,7BB8,7CZ0,7EQ2,7F6U,7FE0,7FE5,7JFT,7JHA,7JHB,7JHC,7JHR,7JHS,7JHT,7JI5,7JI6,7JI7,7JI8,7JI9,7JIM,7JIN,7JIP,7JJ2,7JJ3,7JJ4,7JJ5,7JJ6,7JJU,7JJX,7JJY,7JK0,7JKD,7JKE,7JKG,7JKH,7JKJ,7JKK,7JL9,7JLA,7JLB,7JLD,7JLF,7JNJ,7JNK,7JNL,7JNM,7JOG,7JOI,7JOJ,7JOK,7JOL,7JON,7JP5,7JP6,7JP7,7JP9,7JPB,7JRY,7JRZ,7JS0,7JSB,7JSC,7JVB,7KG1,7KG8,7L2C,7LHK,7MDB,7NRP,7OI1,7Q2U,7QUG,7R3F,7RP2,7RVI,7S6U,7TP3,7XBM,8CYF,8D59,8D5B,8D6T,8EGW,8OEK

Supplementary Table 3 Statistical information on the interactions between the top 10 amino acids with arsenic. Multiple modes were generated for each amino acid's functional groups or side residues, of which only the first ones for each have been mentioned ([a]**Mode**). The statistical parameters such as SE - standard error; SD – standard deviation; Mini – minimum bond length; Max – maximum bond length; Bond – peak bond length and frequency are also shown.

Amino Acids	Side Groups	Mean	SE	Median	[a]Mode	SD	Range(Å)	Mini(Å)	Max(Å)	Bond(Å)	Frequency
Cys	SG	3.05	0.144	3.10	1.90	0.629	2.10	1.90	4.00	2.4	80
	O	3.45	0.104	3.45	2.90	0.361	1.10	2.90	4.00	3.6	08
	N	3.65	0.087	3.65	3.30	0.245	0.70	3.30	4.00	3.5	16
Glu	OE	3.05	0.144	3.10	1.90	0.629	2.10	1.90	4.00	3.6	55
Asp	OD	3.05	0.144	3.10	1.90	0.629	2.10	1.90	4.00	4.0	39
Ser	HA	3.05	0.144	3.10	1.90	0.629	2.10	1.90	4.00	3.0	12
	OG	3.05	0.144	3.10	1.90	0.629	2.10	1.90	4.00	3.8	32
Arg	NH	3.29	0.121	3.30	2.50	0.469	1.50	2.50	4.00	3.7	28
Asn	ND	3.55	0.096	3.55	3.10	0.303	0.90	3.10	4.00	3.8	18
	OD	3.40	0.108	3.40	2.80	0.389	1.20	2.80	4.00	3.5	23
Pro	N	3.75	0.065	3.75	3.60	0.129	0.30	3.60	3.90	3.7	19
His	ND	3.34	0.115	3.35	2.60	0.431	1.40	2.60	4.00	4.0	08
	NE	3.29	0.121	3.30	2.50	0.469	1.50	2.50	4.00	3.8	17
	O	3.80	0.071	3.80	3.60	0.158	0.40	3.60	4.00	3.9	20
Tyr	OH	3.45	0.104	3.45	2.90	0.361	1.10	2.90	4.00	3.8	21
	O	3.45	0.104	3.45	2.90	0.361	1.10	2.90	4.00	3.7	03
Ala	O	3.40	0.100	3.40	2.90	0.332	1.00	2.90	3.90	3.6	20

Supplementary Table 4 Raman bands and assignments for cysteine and its arsenic complex in the spectral region of 150-3000 cm^{-1} in solid and solution form. In this current study, the Raman peaks are assigned based on the literature of Teixeira *et al*., 2007; Zhu *et al*., 2011; Freire *et al*., 2017.

Cys	Cys Solution	CysAsIII Precipitate	CysAsV Precipitate

Raman Shift (cm^{-1})	Assignments Based on Literature	Raman Shift (cm^{-1})	Assignments Based on Literature	Raman Shift (cm^{-1})	Assignments Based on Literature	Raman Shift (cm^{-1})	Assignments Based on Literature
209	CH_2-CH tor	261, 303	CH_2-CH-CO_2 bend	227	CH_2-CH tor	237	CH_2-CH tor
264	NH_2 tor	351	S-H out of the plane bend	285	CH_2-CH-CO_2 bend	266	S-H out of the plane bend
294	CH_2-CH-CO_2 bend	461, 493	H_2O	354, 381, 396, 424	As-S str	300	CH_2-CH-CO_2 bend
360	S-H out of the plane bend	525	CH_2-CH-N bend	496	CH_2-CH-N str	332, 375, 389	As-S str
439	CH_2-CH-CO_2 bend	614	S-S asy str	530, 541	S-S str	480	NH^{3+} tor
536	CH2-CH-N bend	686, 776	C-S str	603	As-O	510	S-S asy
640	CH-CO_2 str	812	CH_2 rock	660	CH-CO2 str	563	NH^{3+} tor
688	C-S str	872	CO_2 sci	670, 692, 709	C-S str	605, 640	CH-CO_2 str
767	CH_2 rock	937	N-CH str	784, 2920	CH_2 rock	670, 688	C-S str
815, 870	CO_2 sci	991	CH_2-CH str	809, 821, 839, 877	CO_2 sci	771	CH_2 rock
937	N-CH str	1063	S-H in-plane bend	918, 937	C-C str	843	CO_2 sci
1003	CH_2-CH str	1099	CH_2 twist	963	N-CH str	997	CH_2-CH
1058, 1064	S-H in plane bend	1141	NH^{3+} rock	1066	S-H plane bend	1056	S-H in-plane bend
1107, 1137	NH^{3+} rock	1212, 1266	CH_2 twist	1104, 1138	NH_3 rock	1205	CH_2 twist
1197	CH_2 twist	1308	C-H bend	1201,1227, 1252	CH_2 twist	1299	C-H bend
1294	C-H bend	1351	CO_2 asy str	1307	C-H bend	1347	CO_2 asy str
1343, 1392	CO_2 asy str	1398, 1428	CH_2 sci	1349	CO_2 asy str	1406	CH_2 sci
1422	CH_2 sci	1637	NH_3 asy bend	1396	C-N str	2548, 2552	S-H str
2550	S-H str	2557	S-H str	1413, 1421	CH_2 sci	2958, 2920	N-H str
2963	N-H str	2954, 3097	N-H str	1497	CH_2/OH bend	2961	CH_2 sym str
2993	C-H str	3001	C-H str	1552, 1599, 2953,	NH_3 asy bend	-	-

Cys	Cys Solution	CysAsIII Precipitate	CysAsV Precipitate
		2964	

Supplementary Table 5 Raman bands and assignments for glutamic acid and its arsenic complex in the spectral region of 150-3000 cm^{-1} in solid and solution form. In this current study, the Raman peaks are assigned based on the literature of Teixeira *et al*., 2007; Zhu *et al*., 2011; Freire *et al*., 2017.

Glu		**Glu Solution**		**GluAsIII Solution**		**GluAsV Solution**	
Raman Shift (cm^{-1})	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments**
244, 280, 315, 385	CCC bend	232	CCC bend	223	CCC bend	225	CCC bend
461	COO^- rock	307	CCC bend	296	CCC bend	300	CCC bend
498	C=O rock	444	C=N bend	357	CCC bend	369	CCC bend
536	C=O out of plane deform	490	C=O rock	413	CCC bend	405	CCC bend
574	COO^- rock	588	C=O bend	486	COO^- bend	487	C=O out of plane bend
706	NH_2 wagg	657	CCC bend	554	C=O out plane bend	516	C=O out of the plane bend
758	COOH wagg	680	NH_2 wagg	590	C=O bend	575	C=O bend
800	CH2 rock/C-COO str/COOH bend	757	NH_2 wagg	702	As-O str	601	C-H out of plane
863	C-C str/CH_2 wagg/COOH bend	793	NH_2 wagg	805	NH_2 rock	640	C=O bend
890, 965, 1062	CC str	854	CC str	855	CH_2 rock	670	NH_2 wagg
916	C-C-N str	899	COOH deform	928	OH bend	738	As-O str
937	COO- deform	915	CCN str	996	CC str	781	CC str
996	NH^{3+} wagg	939	COO- def	1069	CC str	872	As-O str
1084, 1128	NH^{3+} rock	1001, 1026, 1067	CC str	1198	NH_2 twist	921	OH bend
1208	CO str	1097	CN str	1345	CH_2 wagg	970, 1000, 1072	CC str
1248, 1270	CH_2 twist	1204	NH_2 twist	1413	CN str	1192	NH_2 twist

1305	CH_2 wagg	1294	CH_2 twist	1633	H_2O bend	1258, 1287	CH_2 twist
1350, 1433	CH_2 deform	1401	CN str	3149	NH_2 str	1314, 1349	CH_2 wagg
1407	COO^- sym str	1636	H_2O bend			1411	CN str
1482	CH^{3+} deform	3063, 3156	NH_2 str			1634	H_2O bend
2928, 2970	CH_2 str					3137	NH_2 bend

Supplementary Table 6 Raman bands and assignments for aspartic acid and its arsenic complex in the spectral region of 150-3000 cm^{-1} in solid and solution form. In this current study, the Raman peaks are assigned based on the literature of Teixeira *et al*., 2007; Zhu *et al*., 2011; Freire *et al*., 2017.

Asp		**Asp Solution**		**AspAsIII Solution**		**AspAsV Solution**	
Raman Shift (cm^{-1})	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**
187, 244, 266	CCC bend	223, 288	CCC bend	229	CCC bend	208	NH_3 twist
356, 552	CCN bend/CO_2^- bend	483, 599	COOH bend	369	CCN bend/COOH bend	249	CCC bend
412, 462, 776	CO_2^- bend	669	CO_2 bend/HOCC rock/CCN bend	484	OH rock	335	CCN bend/CO_2^- str
597, 658	CO_2^- bend/HOCC rock	989	CC str/CCC bend/CO str	702	As-OH str	369	CCN bend/COOH bend
743	CO_2 bend/ HOCC rock/COOH bend	1059	CN str/CC str	781	OCO^- rock	397	OH rock
872	CC str/CH_2 rock/CO_2^- bend	1080	CN str/CC str	865	CC str/CH_2 rock/CO_2^- bend	511	As-O str
900	CH_2 rock/CC str/CN str	1104	CH str/NH^{3+} rock	984	CC str/CCC bend/CO str	559	As-O str
938	CH str	1150	NH^{3+} rock/CH bend/CO str	1155	NH^{3+} rock/CH bend/CO str	663	CO_2 str/HOCC rock/CCN bend
988	CC str/CCC bend/CO str	1222	CO str/OH bend/CC str	1192	CH str/NH^{3+} rock/CC str/CO str	2926	CH str
1078	CN str/CC str	1356	CC str/CH, OH	1223	CO str/OH	736	CO_2 str/HOCC

			bend/CH_2 wag		bend/CC str		rock/COOH str
1117, 1250, 1332	CH str/NH^{3+} rock	1407	OH bend/CO str/C=O str	1324	CH str/CH_2 wag/NH^{3+} rock	995	CC str/CCC bend/CO str
1141	NH^{3+} rock/CH, CH_2 bend/CO str	1633	H_2O bend	1351	CC str/CH, OH bend/CH_2 wag	1057	CN str/CC str
1286, 1419	C=O str	1073, 1104	CN str/CC str	1642	H_2O bend	1337	CH bend/CH_2 wagg/NH^{3+} rock
1505, 2952, 2995	C=C str, CH_2 str	3064, 3140	NH^{3+} str	3049, 3150	NH^{3+} str	1414	OH bend/CO str/C=O str

Supplementary Table 7 Raman bands and assignments for serine and its arsenic complex in the spectral region of 150-3000 cm^{-1} in solid and solution form. In this current study, the Raman peaks are assigned based on the literature of Teixeira *et al*., 2007; Zhu *et al*., 2011; Freire *et al*., 2017.

Ser		**Ser Solution**		**SerAsIII Solution**		**SerAsV Solution**	
Raman Shift (cm^{-1})	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**
259	CCN bend	297	CCN bend	292	CCC bend	289	CCC bend
291	CCC bend	369	Skel 2	375, 431	Skel 2	383	Skel 2
357	NCC bend	404	Skel 1	491	NH^{3+} twist	422	Skel 1
416	CH_2 bend	491	NH_3^+ twist	591	COO^- wagg	488	NH^{3+} twist
509, 807	COO^- rock	532, 554	COO^- wagg	664	COO^- bend	571	COO^- rock
568		586	COO^- rock	702	As-OH str	654	COO^- bend
606	COO^- bend	593, 602, 673	COO^- bend	807	COO^- rock	709	CH_2 rock
644	CH-CO_2 str	711	CH_2 rock	854	CH_2 rock/CC str	842	CD_2 twist
851, 916	CC str	805	COO^- str	980	CN str	908	CC str
965	CH_2 rock	839	CD_2 twist	1047	CO str	980	CN str
1003	CN str	887	CH_2 rock/CC str	1142	NH^{3+} rock	1058	CO str
1051	CH_2	940	CD_2 wagg	1243	COH bend	1085	NH^{3+} rock
1089	CO str	963, 1030	CO str	1349	C-H bend	1130	NH^{3+} rock

1127	OH bend	989	CN str	1408	COO^- asy str	1157	CH_2 twist
1214	CO str/CH_2 twist	1081, 1111	NH^{3+} rock	1467	CH2 bend	1190	COH bend
1318	CO	1147, 1470	CH_2 twist	1632	NH^{3+} aym bend	1301	CH_2 wagg
1415	COO^- asy str	1241	COH bend	3120	NH^{3+} aym str	1345	CH_2 wagg
1464	CH2 bend	1284, 1348	CH rock			1406	COO^- asy str
1529	CC str	1312	CH_2 wagg			1467	CH_2 bend
1605, 1627	NH^{3+} antisym bend	1402	COO^- sym			1633	HOH bend
1687,1779	C=O str	1634	NH^{3+} antisym bend			3130	NH^{3+} aym str
2958, 2995	CH_2 sym. Str	3038, 3140	NH^{3+} asym str				

Supplementary Table 8 Raman bands and assignments for arginine and its arsenic complex in the spectral region of 300-3000 cm^{-1} in solid and solution form. In this current study, the Raman peaks are assigned based on the literature of Teixeira *et al*., 2007; Zhu *et al*., 2011; Freire *et al*.,2017.

Arg		**Arg Solution**		**ArgAsIII Solution**		**ArgAsV Solution**	
Raman Shift (cm^{-1})	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**	**Raman Shift (cm^{-1})**	**Assignments Based on Literature**
306	CCC bend	343	CCC bend	357	CCC bend	333, 391	As-O str
357	NCC bend	411, 486, 1412	CO_2 bend	399	CO_2 bend	489	CO_2 bend
488	CO_2 bend	597	NCN bend	490, 3167	NH^{3+} twist	596	NCN bend
570, 610	NCN bend	665, 1442	CH_2 wagg	544, 589	NCN bend	681	CH_2 rock
648	CCO bend	781, 804,879	CO_2 /CH_2 rock	578, 653,672	COO^- rock	831	As-OH str
752, 792	CH_2 rock	834, 855,899	CC str/CH_2 rock	702	As-OH str	897	CH_2 rock/CC str
847, 874, 918	CC str/CH_2 rock	981	NCN str/CN asy str/CN str/CC str	730, 789, 858	CO_2 wagg/CH_2 str	930, 1081	NC str/CC str/CH_2 rock/NH_2 wagg
978, 1029	NCN str/CN asy str/CN str/CC str	1058	CC str/NH_2 twist	898	CC str/CH2 rock	979	NCN str/CN sym str/CN str/CC str
1068	CC str/NH_2 twist	1081	NH_2 twist/NH_2 rock/CC str/NC str	930, 1076	NC str/CC str/CH_2 rock/NH_2 wagg	1176	NH_2 wagg/NC str
1096, 1379, 1472	NH_2 twist/NH_2 rock/CC str/NC str	1186	NH_2 wagg/NC str	980	NCN str/CN sym str/CC	1212	NH_3 deform
1121	OH bend	1218	NH^{3+} deform	1136	NH^{3+} rock/CH, CH_2 bend/CO str	1311	CH_2 bend
1189	NH_2 wagg/NC str	1311	CH_2 bend	1173	NH_2 wagg/NC str	1361	COO- motion
1258, 1299	NH^{3+} deform	1358	CH_2 bend/COO- mot	1284, 1409	C=O str	1411	COO- asym str/CH_2 deform
1327, 2863,	CH_2 bend	1470	NCN asy str/CNH	1312, 1361	CH_2 bend	1442	COO- str/CC

2914			bend/ CH_2 wagg				str/NH_2/CH_2 wagg
1432	COO str/CC str/NH_2 twist/CH_2 wagg	1635	CO_2 / NH^{3+} / H_2O	1442	CO_2 str/CC str/NH_2/CH_2 twist	1635	H_2O bend
1553, 1582, 1622	COO- aym str	3149	NH^{3+} str	1635	H_2O bend	3125	NH_3 str