Supporting Information: Stem-loop V of Varkud Satellite RNA Exhibits Characteristics of the Mg^{2+} Bound Structure in the Presence of Monovalent Ions

Christina Bergonzo[1], Kathleen B. Hall[2], Thomas E. Cheatham III[1,]*

[1] Department of Medicinal Chemistry, College of Pharmacy, University of Utah, Salt Lake City, Utah 84112, United States

[2] Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110, United States

* To whom correspondence should be addressed. Tel: 1 (801) 587-9652; Email: tec3@utah.edu

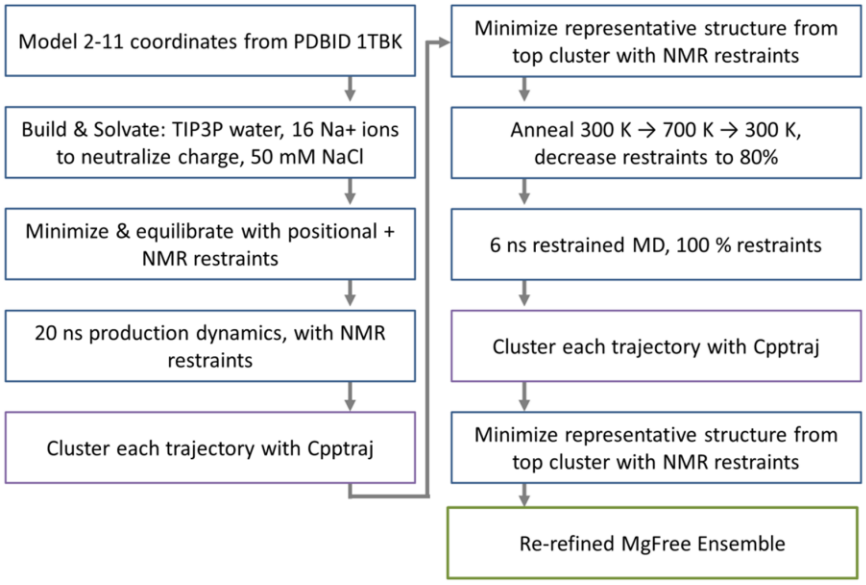
Contents

Supporting Figure 1: Flow chart for MD simulations

1. Initial MgFree Simulations

MgFree + 50 mM NaCl
Model 1 coordinates from PDBID 1TBK
AMBER ff12
TIP3P water, 16 Na^+ ions to neutralize charge, 6 NaCl molecules for [50 mM]
Randomize ion positions x 10
Minimize and equilibrate 10 copies
Production dynamics: 250 ns per copy

2. NMR Re-refinement of MgFree Structure



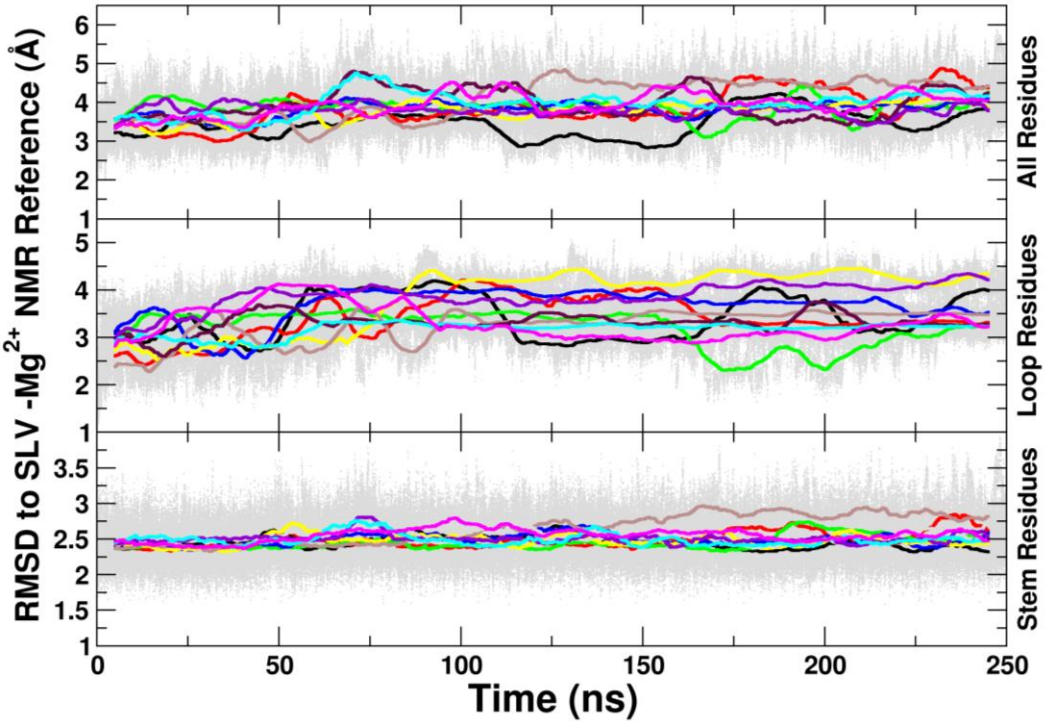
3. Production Dynamics for each SLV RNA (MgFree, MgBound) in each solvent condition (40 mM $MgCl_2$, 50 mM NaCl)

MgBound + 40 mM $MgCl_2$	**MgBound + 50 mM NaCl**
Model 1 RNA coordinates from PDBID 1YN2	
AMBER ff12	
TIP3P water, 8 Mg^{2+} ions to neutralize charge, 5 $MgCl_2$ molecules for [40 mM]	TIP3P water, 16 Na^+ ions to neutralize charge, 6 NaCl molecules for [50 mM]
Randomize ion positions x 10	
Minimize and equilibrate 10 copies	
Production dynamics: 250 ns per copy	

MgFree + 50 mM NaCl	**MgFree + 40 mM $MgCl_2$**
Final minimized coords from new ensemble	
AMBER ff12	
TIP3P water, 16 Na^+ ions to neutralize charge, 6 NaCl molecules for [50 mM]	TIP3P water, 8 Mg^{2+} ions to neutralize charge, 5 $MgCl_2$ molecules for [40 mM]
Randomize ion positions	
Minimize and equilibrate each model	
Production dynamics: 250 ns per model	

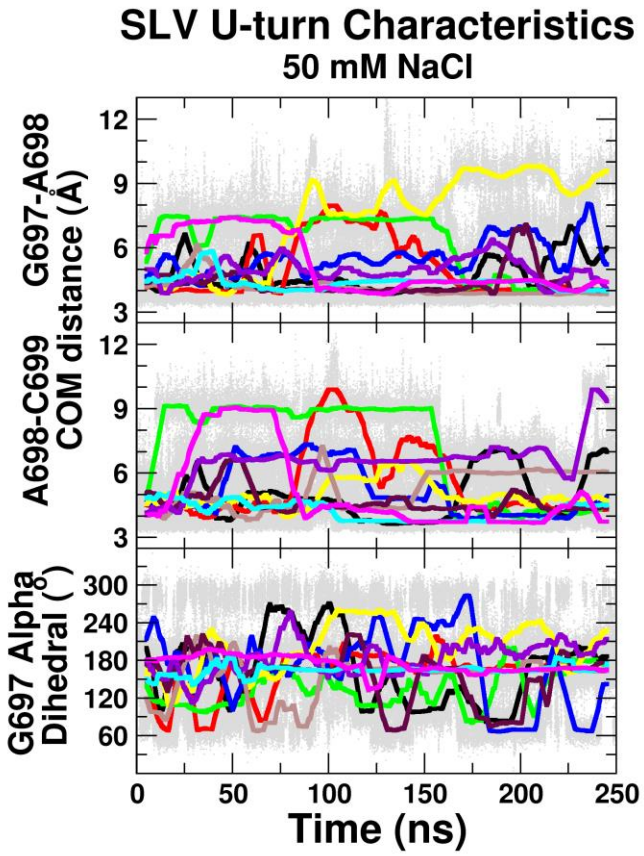
Supporting Figure 2: Heavy atom RMSDs for each of ten independent 250 ns simulations of SLV + 50 mM NaCl, starting from NMR Model 1.

Solid lines are 1000 step running averages. Un-averaged data is shown in grey. Heavy atom RMSDS for All Residues (top), Loop Residues (middle), and Stem Residues (bottom) are shown.



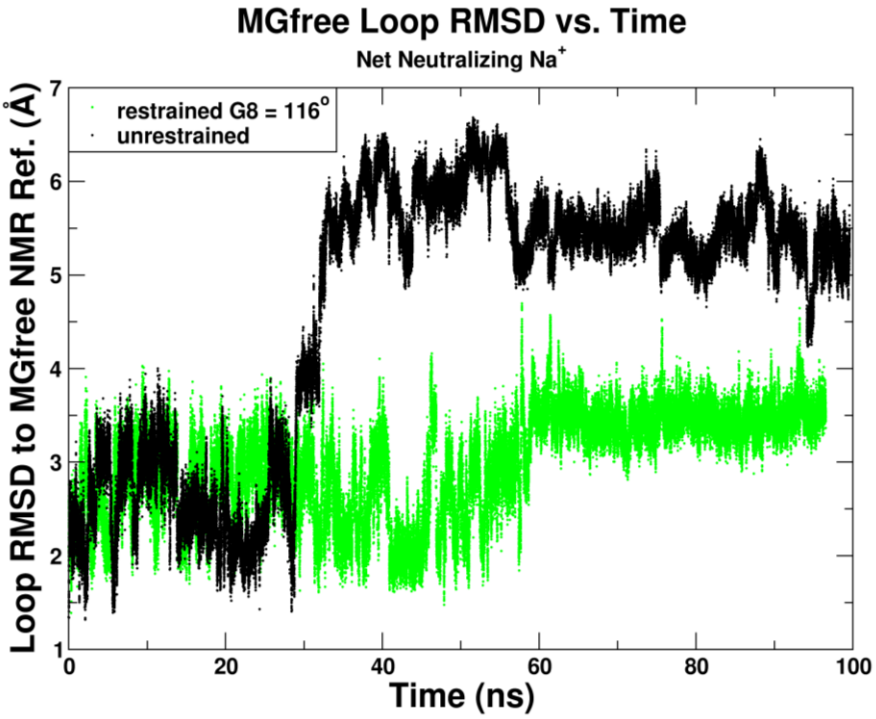
Supporting Figure 3: U-turn Characteristics of SLV in 50 mM NaCl.

Solid lines are 1000 step running averages. Un-averaged data is shown in grey.

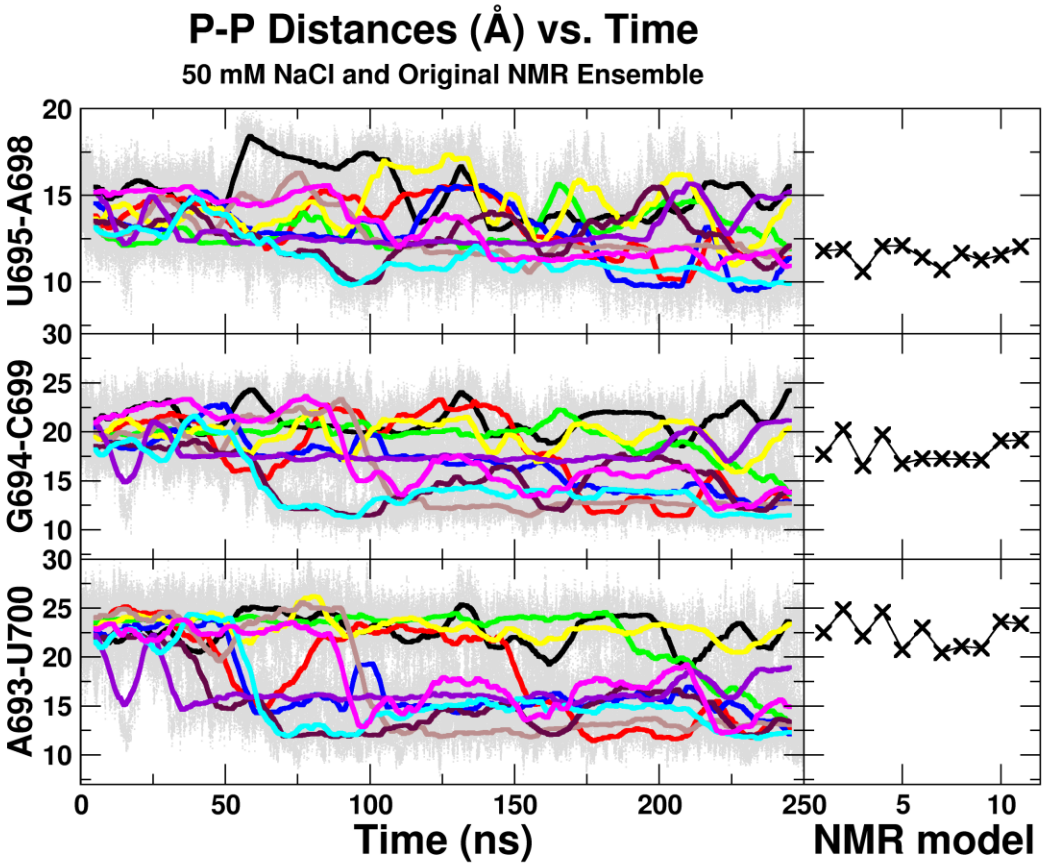


Supporting Figure 4: Loop RMSD to SLV MgFree starting structure.

MgFree starting structure with net neutralizing Na^+ is shown in black. MgFree starting structure with net neutralizing Na^+ and restrained G697 alpha is shown in green.



Supporting Figure 5: Major groove widths during initial MgFree simulations with 50 mM NaCl



Supporting Table 1: Alpha torsion angles in RNA U-turn hairpin loops

System	U-turn	PDB	Loop sequence	alpha	Loop-closing base pair	Method
Varkud SLV –Mg^{2+}	Yes	1TBK	UGACU	116° ±7°	U:A	NMR[b]
Varkud SLV +Mg^{2+}	Yes	1YN2	UGACU	167 ±34°	U:A	NMR[b]
690 loop yeast 16S rRNA	Yes	1FHK	UGAA	172° ±5°	Sheared G:A	NMR[c]
HIV A-rich Stem-loop	Yes	1BJV	UAAA	135° ±99°	Sheared G:A	NMR[d]
Stem-loop IIA U2 snRNA	Yes	2U2A	UAAC	170.6°	Sheared G:A	NMR[e]
Yeast $tRNA^{phe}$	Yes	1EHZ	$A^{OMe}CU^{OMe}GAAWA\psi$[a]	171.1°	C:W	X-ray[f]
Yeast rRNA tetraloop	No	1AFX	UGAA	U-G 300° G-A 55°	C:G	NMR[g]

[a]Abbreviations: ^{OMe}C: 2'O-methyl-C; ^{OMe}G: 2'O-methyl G; W wybutasine; ψ pseudouridine. [b]Campbell and Legault, 2005.[1] *[c]Morosyuk et al., 2001.*[2] *[d]Puglisi and Puglisi, 1998.*[3] *[e]Stallings and Moore, 1997.*[4] *[f]Shi and Moore, 2000.*[5] *[g]Butcher et al., 1999.*[6]

(1) Campbell, D. O.; Legault, P. Nuclear Magnetic Resonance Structure of the Varkud Satellite Ribozyme Stem-Loop V RNA and Magnesium-Ion Binding from Chemical-Shift Mapping. *Biochemistry* **2005**, *44*, 4157–4170.

(2) Morosyuk, S. V; Cunningham, P. R.; SantaLucia, J. Structure and Function of the Conserved 690 Hairpin in Escherichia Coli 16 S Ribosomal RNA. II. NMR Solution Structure. *J. Mol. Biol.* **2001**, *307*, 197–211.

(3) Puglisi, E. V.; Puglisi, J. D. HIV-1 A-Rich RNA Loop Mimics the tRNA Anticodon Structure. *Nat. Struct. Biol.* **1998**, *5*, 1033–1036.

(4) Stallings, S. C.; Moore, P. B. The Structure of an Essential Splicing Element: Stem Loop IIa from Yeast U2 snRNA. *Structure* **1997**, *5*, 1173–1185.

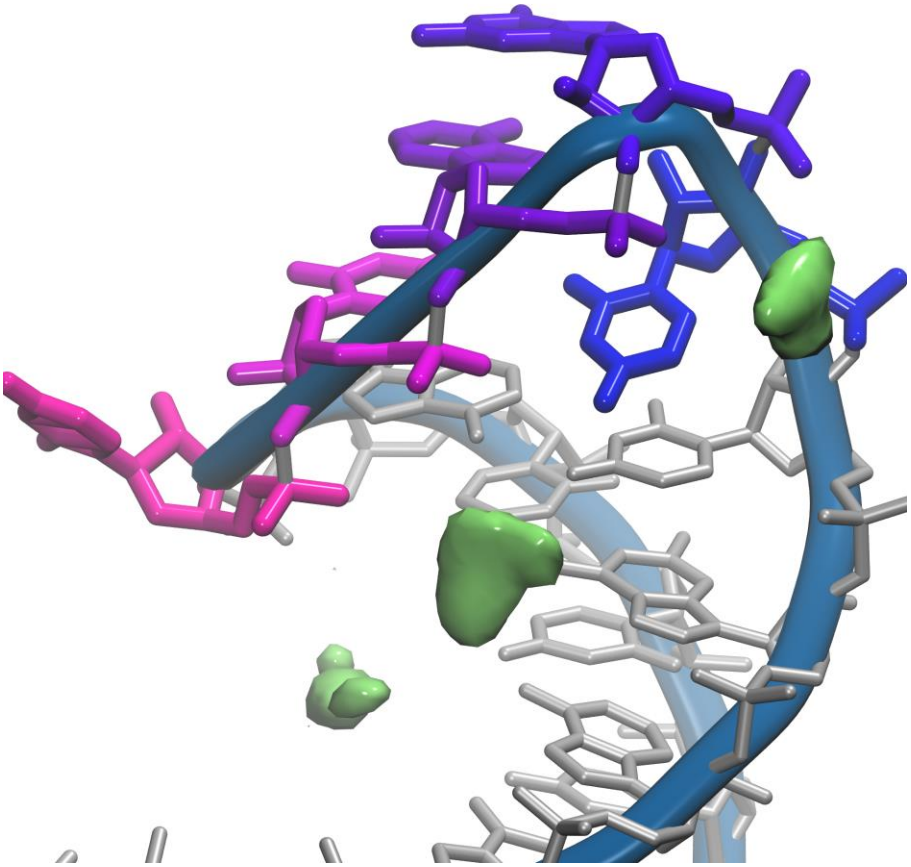
(5) Shi, H.; Moore, P. B. The Crystal Structure of Yeast Phenylalanine tRNA at 1.93 A Resolution: A Classic Structure Revisited. *RNA* **2000**, *6*, 1091–1105.

(6) Butcher, S. E.; Allain, F. H.; Feigon, J. Solution Structure of the Loop B Domain from the Hairpin Ribozyme. *Nat. Struct. Biol.* **1999**, *6*, 212–216.

Supporting Table 2: Clusters' r^6 averaged values for MgBound NOE violations

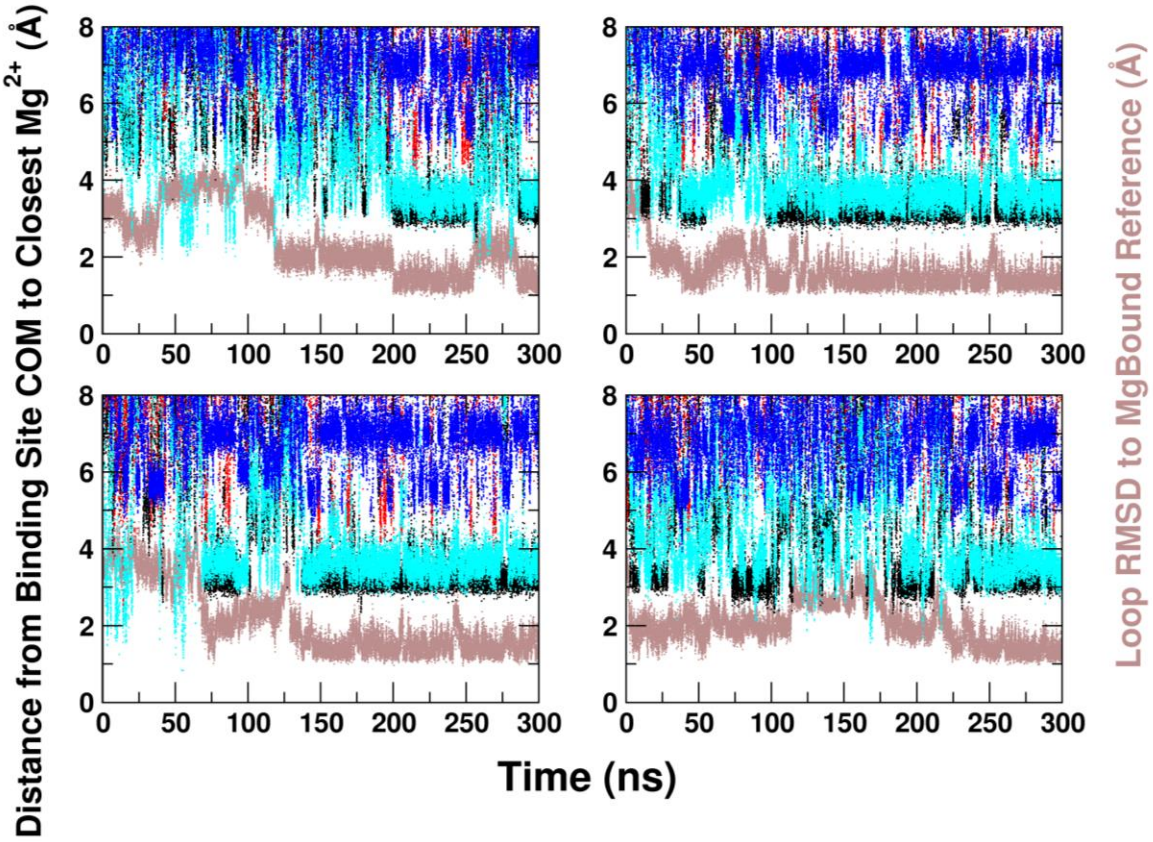
	NOE	Divalent Ions, Top Cluster	Monovalent Ions, Top Cluster	Divalent Ions, 2nd Cluster
Strong NOEs	**:699@H4'-:699@H3'**	3.0164	2.9966	3.0108
1.8-3	**:697@H4'-:697@H3'**	3.0053	2.997	2.9983
	:696@H4'-:696@H3'	3.0147	2.9999	3.0068
	:698@H3'-:698@H8	2.9353	3.4673	2.9153
	:696@H1'-:699@H5	2.7024	4.458	4.2819
	:696@H2'-:698@H8	2.6593	3.2372	2.9893
Medium NOEs	**:696@H1'-:701@H2**	4.1875	5.0432	4.9356
1.8-4.1	**:699@H2'-:700@H6**	4.1788	4.98	4.4771
	:700@H3'-:700@H6	4.4508	3.9206	4.3991
	:698@H2-:699@H1'	3.3934	3.541	4.3702
	:697@H2'-:697@H8	4.1629	4.2032	4.0412
Weak NOEs	**:700@H5'-:701@H8**	5.8029	5.4213	5.6661
1.8-5.5	**:699@H3'-:700@H5**	6.4325	6.6634	6.5658
	:699@H3'-:700@H6	5.5373	5.5862	5.5466
	:696@H2'-:699@H5	4.5356	5.657	5.8133
	:696@H2'-:698@H3'	4.678	5.7503	4.9862
	:695@H4'-:696@H6	5.5079	6.1696	5.9121
	:695@H1'-:696@H6	5.3343	5.7325	5.3803
V. Weak NOEs	**:698@H5'-:699@H5**	7.7272	4.1343	6.4342
1.8-7.0	**:696@H4'-:699@H5**	5.521	7.4328	7.3851

Supporting Figure 6: Localization of top 10% of Mg^{2+} ion density (shown in green) of 2nd most populated cluster in the divalent ion environment simulations



Supporting Figure 7: Closest Mg^{2+} to Binding Sites 1-4 in MgFree + 40 mM $MgCl_2$ Simulations

Four out of ten MgFree + 40 mM $MgCl_2$ simulations sample the MgBound conformation. Below is the time dependent plots of closest Mg^{2+} ion to the centers of mass of binding sites 1 (black), 2 (red), 3 (cyan), and 4 (blue) for each of these four simulations. The loop RMSD to the MgBound reference is shown for each simulation in brown. The scales for distance and RMSD are the same.



Supporting Table 3: Percent Occupancy of atoms designated as Binding Sites 1-4 from Campbell et al. 2006

Binding Site	Residue	Atom	Divalent Ions, Top Cluster		Divalent Ions, 2nd Cluster	
			Occupancy (%)	Avg Dist (Å)	Occupancy (%)	Avg Dist (Å)
1	U_{696}	OP1	1.52	3.84	7.32	3.83
	U_{696}	OP2	43.35	3.75	15.13	3.80
	U_{696}	O5'	0.01	3.96	0.08	3.93
	G_{697}	OP1	3.05	3.84	2.91	3.85
	G_{697}	O3'	0.01	3.92	0.03	3.94
	A_{698}	OP1	7.85	3.87	6.60	3.85
	A_{698}	OP2	32.11	3.87	18.6	3.85
2	U_{696}	O2'	0.05	3.88	0.03	3.93
	G_{697}	N7	0.62	3.87	0.38	3.87
	A_{698}	N7	---	---	---	---
3	U_{695}	O4	16.92	3.86	20.45	3.83
	U_{696}	O4	5.67	3.84	7.32	3.83
	C_{699}	OP1	14.13	3.88	13.16	3.85
	C_{699}	OP2	23.11	3.86	9.17	3.87
4	U_{700}	OP1	7.78	3.87	5.48	3.86
	U_{700}	O3'	0.05	3.92	0.05	3.93
	A_{701}	N7	---	---	---	---

Supporting Script 1: Clustering command for analysis of restrained re-refined SLV simulations.

```
parm nowat.50mMNaCl.parm7
trajin nowat.nc
cluster c01 :1-17&!@H epsilon 1.0 sieve 5 out cnumvtime.dat
summary summary.dat info info.dat
```

Supporting Script 2: Input file for simulated annealing.

```
Production run
 &cntrl
    ntpr = 500, ntwr = 500, ntwx = 500, ntwe = 500,
    ntf = 2, ntc = 2, ntb = 1, ntp = 0, nscm = 500,
    ntt = 3, ig=-1, gamma_ln=1.0,
    tautp = 0.5, taup = 5.0, cut = 9.0,
    nstlim = 2000000, dt = 0.002, iwrap = 1,
    irest = 0, ntx = 1, tol = 0.00000001,
    ioutfm=1,
    nmropt=1, pencut=-1, tempi=300.0,
 /
 &wt type = 'REST',
istep1=0,istep2=750000,value1=0.08,value2=0.08, /
 &wt type = 'REST',
istep1=750001,istep2=1000000,value1=0.08,value2=1.0, /
 &wt type = 'REST',
istep1=1000001,istep2=2000000,value1=1.0,value2=1.0, /
 &wt type='TEMP0', istep1=0,    istep2=500000,  value1=300.0,
value2=700.0, /
 &wt type='TEMP0', istep1=500001, istep2=1000000, value1=700.0,
value2=700.0, /
 &wt type='TEMP0', istep1=1000001, istep2=2000000, value1=700.0,
value2=300.0, /
 &wt type = 'END', /
   DISANG=../all.RST
   LISTOUT=POUT
 &wt type='END', &end
END
```

Supporting Script 3: Cpptraj command to separate MgBound and MgFree SLV simulation data < 2.5 Å RMSD to MgBound loop.

```
cpptraj << EOF
parm nowat.rna-ions.parm7
```

```
trajin nowat.rna-ions.nc
autoimage
readdata rmsd-1yn2-loop.dat name data
outtraj rmsmax_2.5.nc maxmin data:2 min 0.0 max 2.5
go
EOF
```

Supporting Script 4: Clustering command for analysis of MgBound and MgFree SLV simulation data < 2.5 Å RMSD to MgBound loop.

```
cpptraj <<EOF
parm nowat.rna-ions.parm7
trajin rmsmax_2.5.nc
cluster kmeans clusters 2 rms mass :6-12@C*,N*,O*,P* sieve 10
out cvt.dat summary summary.dat clusterout cluster.nc clusterfmt
cdf repout rep repfmt pdb
EOF
```

Supporting Script 5: Grid density analysis.

```
cpptraj << EOF
parm nowat.rna-ions.parm7
trajin cluster.c0.nc
autoimage origin
parm nowat.rna.parm7
reference average.c0.MGfree.noions.pdb parm nowat.rna.parm7
rms reference :1-17
#Calculate the histogram of MG ions in a 50x50x50Angstrom  box,
with a grid spacing of 0.5Angstroms in x,y,z
#Isovalue is incremented every time a MG is within a grid space
#Density is calculated as number of ions/volume
grid grid_Na.c0.D.dx 100 0.5 100 0.5 100 0.5 :MG normdensity
density 0.00007135
EOF
```

Supporting Script 6: Hbond analysis.

```
cpptraj << EOF
parm nowat.rna-ions.parm7
trajin cluster.c0.nc
hbond h1 dist 4.5 :1-17 solventdonor :Na+ solventacceptor :Na+
out allUV.dat avgout avg.dat solvout solv.dat bridgeout
bridge.dat
EOF
```